

RESULT 5

US-09-056-556-179
Sequence 179, Application US/09056556
Patent No. 6350456

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

TREATM

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 179:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-056-556-179

Query Match 3.9%; Score 49.6; DB 4; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.0015;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY	458	CCTCAGTGGAGGTTGCGGAGGCGAGCAACCCAGAGGAGGAGGAGCCCGGCAT	517
DB	119	CTTCAACCAAGGCGGCGAGCGCAACCGCGCAACCGCGGTGACGGCGGGGTGCGGCGCA	178
QY	518	CCTCGGGAGCCCGGGAACCGGAGCGGAGAGGAGGAGGAGGAGCCCGGCAT	577
DB	179	CGGCGGAAACGGCGGAAACGGCGGCAACACACACCGCCGCGCGGCGGCAACGAG	238
QY	578	CGGAAACGAGAGGCGGATGTTCAAGCTGCGAGGCGAGCGCGGTGGAGCAGCCAC	637
DB	239	CGGCGAGCGGCGGCGCGCGCGCGCGCGGAGACCGCGGAGCGCGCGCGGCAC	298
QY	638	CACCGAGGAGCGGAGACAGTGTGCGGCGAGTGAGTCTGAGCCTGGAGACGAGCCAT	697
DB	299	CGGACCGGCGGCGGCAACAGGCAACGGCGGCAACGGCGGCGGCGGCAAGGCGGCAC	358
QY	698	GTCCTGCCGAGAGATGCTCAGG	721
DB	359	CGGCGGAGCGGTGCACTTCAGG	382

RESULT 6

US-09-072-596-174
Sequence 174, Application US/09072596
Patent No. 6458366
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, David C.
APPLICANT: Campos-Neto, Antonia
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.

APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 174:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-072-596-174

Query Match 3.9%; Score 49.6; DB 4; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.0015;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY	458	CCTCAGTGGAGGTTGCGGAGGCGAGCAACCCAGAGGAGGAGGAGCCCGGCAT	517
DB	119	CTTCAACCAAGGCGGCGAGCGCAACCGCGCAACCGCGGTGACGGCGGGGTGCGGCGCA	178
QY	518	CCTCGGGAGCCCGGGAACCGGAGCGGAGAGGAGGAGGAGGAGGAGCCCGGCAT	577
DB	179	CGGCGGAAACGGCGGAAACGGCGGCAACACACACCGCCGCGCGGCGGCAACGAG	238
QY	578	CGGAAACGAGAGGCGGATGTTCAAGCTGCGAGGCGAGCGCGGTGGAGCAGCCAC	637
DB	239	CGGCGAGCGGCGGCGCGCGCGCGCGCGGAGACCGCGGAGCGCGCGCGGCAC	298
QY	638	CACCGAGGAGCGGAGACAGTGTGCGGCGAGTGAGTCTGAGCCTGGAGACGAGCCAT	697
DB	299	CGGACCGGCGGCGGCAACAGGCAACGGCGGCAACGGCGGCGGCGGCAAGGCGGCAC	358
QY	698	GTCCTGCCGAGAGATGCTCAGG	721
DB	359	CGGCGGAGCGGTGCACTTCAGG	382

RESULT 7

US-09-072-967-179
Sequence 179, Application US/09072967
Patent No. 6592877
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Devin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mark, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121,411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 179:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-967-179

Query Match 3.9%; Score 49.6; DB 4; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.0015;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 458 CCTACTGAGGTTGGCGGAGGCGAGGACACCCAGAGGAGGAGGAGCCCGGCAT 517
DB 119 CTTACCCCAAGCGCGCGCAACGCGGCAACGCGGTGACGCGCGGTGGCGGCA 178
QY 518 CCTCGGGGACCCGGGAAACCGAGCGGAGAGGTGGCCGCGAGGAAACGCGTCCCCAGC 577
DB 179 CGGCGGAAACGCGGAAACGCGGCAACACACACACGCGCGCGCGCGCACACAGG 238
QY 578 CGCGAAGGAGAGGCGGATTTCAAGGTGAGGCGGAGCGCGGTGGAGACAGCAGCAC 637
DB 239 CGGCGACGCGCGGCGCGCGCGCGGCGGAAACCGCGGAAACCGCGACCGCGGCAC 298
QY 638 CACCGAGAGACCGGAACAGTGTGGCGGATGAGTCTGAGCTGAGAGGAGAGGACAT 697
DB 299 CGGACACGCGCGGCAACAGGCAACGCGGCAACGCGGCGGACCGCGCGCAAAAGCGGCAC 358
QY 698 GTCTGCGCGAGAGTGCTCAGG 721
DB 359 CGGCGGCGACGCTGACTCTCAGG 382

RESULT 8
US-08-455-543A-8
Sequence 8, Application US/08455543A
Patent No. 5792846
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven

APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: May 31, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-455-543A-8

Query Match 3.8%; Score 49; DB 1; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGCTGAGTGAACCCGCGGAGCGCTCGTGAAGTGTCACTCGAGCGC 309
 Db 2751 GAGGCCCCGAGAGGCGAGAGCGGAGAGCCCGTGGCCCGGAGAGCGCGCGCGAC 2810
 QY 310 CGCAGCTGAGCTGAGAGGTGCGAGCGAGCGCCCGAGTCGGGTTCCTGTGA 369
 Db 2811 CGCA-GCCACAGAGAGAGAGCGCGGAGCC--CCGAGAGCGCGAGCGCGCGCA 2867
 QY 370 CCCGTTGCGCAGCGCGCGCGAGATCCCGCGATCTTGCGAGACCGTACCGCGTTC 429
 Db 2868 GGCCTCAGCGCCGAGGCGCGCGCGCGCAACCGCGCGCGTCCCGAGAGCGCGCC 2927
 QY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCACTGAGAGTTCGGAGAGCGAGCA 489
 Db 2928 GAGCGGAGAGCGCGAGCGCAACCGCGAGCGCGAGCGAGAGAGGCGCC 2987
 QY 490 CCCACGAGAGAGAGAGAGCGCGCGATCTCGGAGACCGGAGCGAGAGAG 549
 Db 2988 GGGCGCAAGGAGAGAGCGCGCGCGCGAGCACCGCGCGCGCGCGCGCGAG 3047
 QY 550 GTGGCGCGGAGAGAGCGGTCCCGCAGCGCGAGAGAGAGGCGATGTTCAAGCTGCA 609
 Db 3048 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCACCGCGCGCGCGCACAGCGCGCTGCT 3107
 QY 610 GGGCAGAGCGGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 668
 Db 3108 CACGAGGCTGTGAGAGAGAGAGCGAGAGAGAGCGCGCGCGCGCGCGCGAT 3166

RESULT 9
 US-08-193-078B-8
 ; Sequence 8, Application US/08193078B
 ; Patent No. 5846757
 ; GENERAL INFORMATION:
 ; APPLICANT: Harpold, Michael
 ; APPLICANT: Ellis, Steven
 ; APPLICANT: Williams, Mark
 ; APPLICANT: Feldman, Daniel
 ; APPLICANT: McCue, Ann
 ; APPLICANT: Brenner, Robert
 ; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
 ; STREET: 1660 UNION STREET
 ; CITY: SAN DIEGO
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/193,078B
 ; FILING DATE: 07-FEB-1994
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/868,354
 ; FILING DATE: 10-APR-1992
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/745,206
 ; FILING DATE: 15-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 6362-53607
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-238-0999

TELEFAX: 619-238-0062
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7175 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 144..6857
 ; FEATURE:
 ; NAME/KEY: 5' UTR
 ; LOCATION: 1..143
 ; FEATURE:
 ; NAME/KEY: 3' UTR
 ; LOCATION: 6855..7175
 ; US-08-193-078B-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
 Best Local Similarity 48.0%; Pred. No. 0.0085;
 Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGCTGAGTGAACCCGCGGAGCGCTCGTGAAGTGTCACTCGAGCGC 309
 Db 2751 GAGGCCCCGAGAGGCGAGAGCGGAGAGCCCGTGGCCCGGAGAGCGCGCGAC 2810
 QY 310 CGCAGCTGAGCTGAGAGGTGCGAGCGAGCGCCCGAGTCGGGTTCCTGTGA 369
 Db 2811 CGCA-GCCACAGAGAGAGAGCGCGGAGCC--CCGAGAGCGCGAGCGCGCGCA 2867
 QY 370 CCCGTTGCGCAGCGCGCGCGAGATCCCGCGATCTTGCGAGACCGTACCGCGTTC 429
 Db 2868 GGCCTCAGCGCCGAGGCGCGCGCGCGCAACCGCGCGCGTCCCGAGAGCGCGCC 2927
 QY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCACTGAGAGTTCGGAGAGCGAGCA 489
 Db 2928 GAGCGGAGAGCGCGAGAGCGCGCGCGAGCACCGCGCGCGCGCGCGAG 3047
 QY 490 CCCACGAGAGAGAGAGAGCGCGCGATCTCGGAGACCGGAGCGAGAGAG 549
 Db 2988 GGGCGCAAGGAGAGAGCGCGCGCGCGCGCACCGCGCGCGCGCGCGCGAG 3047
 QY 550 GTGGCGCGGAGAGAGCGGTCCCGCAGCGCGAGAGAGAGGCGATGTTCAAGCTGCA 609
 Db 3048 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCACCGCGCGCGCGCGCGAGCTGCT 3107
 QY 610 GGGCAGAGCGGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 668
 Db 3108 CACGAGGCTGTGAGAGAGAGAGCGAGAGAGAGCGCGCGCGCGCGCGCGAT 3166

RESULT 10
 US-08-223-305C-8
 ; Sequence 8, Application US/08223305C
 ; Patent No. 5851824
 ; GENERAL INFORMATION:
 ; APPLICANT: Harpold, Michael
 ; APPLICANT: Ellis, Steven
 ; APPLICANT: Williams, Mark
 ; APPLICANT: Feldman, Daniel
 ; APPLICANT: McCue, Ann
 ; APPLICANT: Brenner, Robert
 ; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brown, Martin, Haller & McClain
 ; STREET: 1660 Union Street
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92101-2926

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-223-305C-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTCGGGTGAGGAGCCGCGCGGAGCGCTCGGTGAGGTTCACCTCGGGCGC 309
DB 2751 GAGGCCCCGAAAGCGGAGAGGGGAGCCCGGTGCGGAGAGCGCGCCGCGCAC 2810
QY 310 CGCAGCTGCGAGCTGCGAGGTGCGGAGCCGCGCGGATCGGCTTCTGTCAA 369
DB 2811 CGCA-GCCAGAGCAAGAGAGGCGCGGGCC--CCCGAGCGCGGAGCGCGCGCGA 2867
QY 370 CCCCGTGGCCACGCGCGCGCGGAGATCCCCCGGATCTGTGAGACCGTACGCCCTTC 429
DB 2868 GGGCCAGGCGCGCGGCGCGCGCGGACACCGCGCGCTCCCGGAGGAGGCGGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCACTGAGAGTTGCGGAGGCAAGACAGCA 489
DB 2928 GAGCGGAGAGCCCGAGCGCAACCGCGCGGACCGGACCAAGATCCAGCAAGAGTGGCC 2987

QY 490 CCCAGAAAGGAGAGGAGAGCCCGGCACTCTCGGGAACCCGGAAACCGGAGAGAG 549
DB 2988 GAGCGCAAGGGCGAGCGCGCGCGGCAACCGCGCGGCGCGCGCGCGCGAG 3047
QY 550 GTGGCGCGGAGAAAGCGGTCCCGGAGCGGAGAGGCGGATGTTCAGGCTGCA 609
DB 3048 GCGAGAGCGGGGAGAGAGCGCGCGCGCGGACCGCGCGCGCGCGCGAGCTGCT 3107
QY 610 GGGCAGCGCGGTGAGAGCAGACCAACCGAGAGACCGGAAACAGTGTGCGCGCAT 668
DB 3108 CACGAGGCTGTGAGAGAGAGCAGCAGAGAGAGGCGCACGAGAGAGGCTGAGAT 3166

RESULT 11
US-08-149-097D-8
Sequence 8, Application US/08149097D
Patent No. 5874236
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,097D
FILING DATE: 05-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,231
FILING DATE: 13-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899

FILING DATE: 04-APR-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 6362-55038
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 238-0999
 TELEFAX: (619) 238-0062
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7175 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 144..6857
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..143
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 6855..7175
 US-08-149-097D-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
 Best Local Similarity 48.0%; Pred. No. 0.0085;
 Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

250 GACGCTGCGGTGAGTGAACCCGCGCCGCGACGCTTGGTGAAGTTCACCTCGGCGC 309
 2751 GAGGCCCCCAAGCGGAGAGCGGAGCGCGGCGCGGCGCGCGCGCGCGCGCGCGC 2810
 310 CGGAGCGCTGAGCGCTGCGAGGTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGC 369
 2811 CGCA-GCCACAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2867
 370 CCCCGTGGCCAGCGCGCGCGCGCGAGATCCCGCGATCTTGCGACACCGTACCGCTTC 429
 2868 GCGCCAGC 2927
 430 TCGTCGTGACCTTCTGTGAGCTTCTCTCTCACTGAGGTTGCGGAGCGCAGACACA 489
 2928 GAGCGGGAACCCCGACCGCACCGCGCGCACCGGACACAGATCCAGCAAGAGTGCGCC 2987
 490 CCACAGAGAGAGAGAGAGCGCGCGCGCGATCTCGGAGAACCGGAGCCGAGCGAGAG 549
 2988 GCGCGCAAGCGGAGAGC 3047
 550 GTGGCCCGCAGAGAAAGCGGTCCCGACGCGCGAGCGGAGCGGATGTTCAAGCTGCA 609
 3048 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCGCACCGGCGCGCGCAAGAGCGCGCTGCT 3107
 610 GGGCAGAGCGGAGGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 668
 3108 CACGAGGCTGTGAGAGAGACCAAGAGAAAGAGCGCAGAGAGAGAGGCTGAGAT 3166

RESULT 12
 US-08-949-386-8
 Sequence 8, Application US/08949386
 Patent No. 6090623

GENERAL INFORMATION:
 APPLICANT: Harpold, Michael
 APPLICANT: Williams, Steven
 APPLICANT: Williams, Mark
 APPLICANT: McCue, Ann
 APPLICANT: Gillespie, Alison
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 METHODS
 NUMBER OF SEQUENCES: 38
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McCain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: US
 ZIP: 92101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/949,386
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,012
 FILING DATE: 11-AUG-1994
 APPLICATION NUMBER: 08/149,097
 FILING DATE: 5-NOV-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/105,536
 FILING DATE: 11-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 519808
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 238-0999
 TELEFAX: (619) 238-0062
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7175 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 144..6857
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..143
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 6855..7175
 US-08-949-386-8

Query Match 3.8%; Score 49; DB 3; Length 7175;
 Best Local Similarity 48.0%; Pred. No. 0.0085;
 Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

250 GACGCTGCGGTGAGTGAACCCGCGCCGCGACGCTTGGTGAAGTTCACCTCGGCGC 309
 2751 GAGGCCCCCAAGCGGAGAGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGC 2810
 310 CGGAGCGCTGAGCGCTGCGAGGTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGC 369
 2811 CGCA-GCCACAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2867
 370 CCCCGTGGCCAGCGCGCGCGCGAGATCCCGCGATCTTGCGACACCGTACCGCTTC 429
 2868 GCGCCAGC 2927
 430 TCGTCGTGACCTTCTGTGAGCTTCTCTCTCACTGAGGTTGCGGAGCGCAGACACA 489
 2928 GAGCGGGAACCCCGACCGCACCGCGCGCACCGGACACAGATCCAGCAAGAGTGCGCC 2987
 490 CCACAGAGAGAGAGAGAGCGCGCGCGCGATCTCGGAGAACCGGAGCCGAGCGAGAG 549
 2988 GCGCGCAAGCGGAGAGC 3047
 550 GTGGCCCGCAGAGAAAGCGGTCCCGACGCGCGAGCGGAGCGGATGTTCAAGCTGCA 609

Db 3108 CACGAGGCTGTGAGAGAGAGACCAACGAGAGAGAGGCGGACCGAGAGAGAGGCTGAGAT 3166

RESULT 14
US-08-984-709A-8
Sequence 8, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Stauderman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,709A
FILING DATE: 02-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 587-5360
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-984-709A-8

Query Match 3.8%; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

Db 250 GACGCTGGCGTGCAGGTGAACCCGCGCGGACGCGCTGCGTGCAGTGTCTACTCGGGGCC 309
2751 GAGGCCCCGAGAGGCGGAGAGCGCGGAGCGCCGGTCCCGGAGAGCGGCGCGCGCAC 2810
310 CGCAGCGTGCAGCTGCAGAGGTGCGGAGCGGACCGCGCGCGATCGAGTTCCTGTCAA 369
2811 CGCA-CCGACAGCAAGAGAGCGCGCGGCGC--CCCGAGAGCGCGGAGAGCGCGCGCGCA 2867
370 CCCGCTGCGCAGCGCGCGCGCGGAGATCCCGCGGATCTTGCAGACCGTAGCCCCGCTTC 429
2868 GCGCCAGCGCGCGAGGCGCGCGCGGAGCAGCAGCGCGCGCGCTCCCGAGAGAGCGCGCC 2927
430 TCGTCGAGACTTGTGCGCTCTCTCTCTCACTGAGAGTGTGGGAGAGGAGGACAGACA 489

Db 2928 GAGCGGAGCGCCCGACCCGACCGCGCGCACCGGACCGGACCGAGATCCGAGCAAGAGTGGCC 2987
490 CCCAGAGAGAGAGGAGGAGCGCGCGCATCCTCGGAGACCGCGGAAACCGAGCGGAGAG 549
2988 GCGCCAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047
550 GTGCGCGCGAGAGAAAGCGGTCCCGCAGCGCGGAGCGAGGCGGATGTTCAAGCTGCA 609
3048 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3107
610 GCGGAGCGCGCGTGGAGAGCAGCAGCAGCAGCGGAGAGCGGAAAGTGTGGCGCGAT 668
Db 3108 CACGAGGCTGTGAGAGAGAGACCAACGAGAGAGGCGGACCGAGAGAGGCTGAGAT 3166

RESULT 15
US-08-450-272-8
Sequence 8, Application US/08450272
Patent No. 6387696
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Alison
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,272
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,350
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: 08/336,257
FILING DATE: 7-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083
FILING DATE: 28-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/311,363
FILING DATE: 23-SEPT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,012
FILING DATE: 11-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: 4-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,078
FILING DATE: 07-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,097

FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
TELEPHONE: (619) 238-0999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-272-8

Query Match 3.8%; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGGTGCAAGTGAACCGCGCGGACGCTCGGTGCACTGCGGCGC 309
DB 2751 GAGGCCCCGGAAGCGGAGCGGAGCGCGGTGCGCGGAGGCGCGCGCGCAC 2810
QY 310 CGCAGCGTGGAGCTGAGGGGTGCCGAGCCGAGCCGAGTCCGGTTCGTCTCA 369
DB 2811 CGCA-CGCAAGCAAGAGGCGCGCGGCTC-CCGAGGCGGAGCGAGCGCGCGCA 2867
QY 370 CCCGTGGCCACCGCGCGCGGAGATCCCGCGCATCTTGCGCAGACCGTAAGCCCGTTTC 429
DB 2868 GAGCCAGGCGCGGAGGCGCGCGGCGCACACCGCGCGGCTCCCGAGAGGAGCGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAAGTTGCGGAGGCAAGCAGACA 489
DB 2928 GAGCGGAGCGCGGAGCGCGCACCGCGCGCACCGGACCAAGATCCGAGCAAGAGTGGCC 2987
QY 490 CCCAGCAAGGAGAGGAGCGGAGCCCGGATCCTCGGAGACCGGAGCCGAGAGAG 549
DB 2988 GCGCGCAAGGAGGAGGCGCGCGCGGCAACCGCGCGGCTCCCGAGCGGCGCGCGAG 3047
QY 550 GTGGCCGCGAGAAAGCGGTCCCGCAGCCGCGCAAGCAGAGAGGCGATGTTCAAGCTGCA 609
DB 3048 GCGGAGAGGCGGAGGAGCGCGCGCGCGCACCGGCGCGCGCACCAAGCGCGAGCTGCT 3107

QY 610 GGGCAGCGCGGTGGAGAGCAGCACCACCGGAGACCGGAACAGTGTGCGCGCAT 668
DB 3108 CACGAGCTGTGAGAGGAGGACCAAGAGAAAGAGGCCACGAGAGAGAGGCTGAGAT 3166

Search completed: April 6, 2004, 20:38:15
Job time: 138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 16:19:29 / Search time 3686 Seconds

(Without alignments)
10345.633 Million cell updates/sec

Title: US-09-830-810a-1
Perfect score: 1277
Sequence: 1 aagcgcgagcgagcgagcgagga.....acaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: EST.*
1: em_estba.*
2: em_estbm.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gse_hum.*
18: em_gse_inv.*
19: em_gse_pin.*
20: em_gse_vrt.*
21: em_gse_fun.*
22: em_gse_mam.*
23: em_gse_mus.*
24: em_gse_pro.*
25: em_gse_xod.*
26: em_gse_pig.*
27: em_gse_vr1.*
28: gb_gse1.*
29: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	643	50.4	675	10	BB641267
2	642.4	50.3	666	10	BB264222
3	591.4	46.3	615	10	BB071693
4	582.4	45.6	595	12	BB084538

C	5	570.4	44.7	594	10	BF471866	BF471866	UI-M-BH3-
C	6	497.2	38.9	525	10	BB704019	BB704019	BB704019
C	7	473.8	37.1	499	9	AT854700	AT854700	UI-M-BH0-
C	8	448.4	35.1	491	10	BB703259	BB703259	BB703259
C	9	413.8	32.4	491	10	BB704648	BB704648	BB704648
C	10	405.8	31.8	436	10	AU044294	AU044294	AU044294
C	11	398.2	31.2	425	10	BB705931	BB705931	BB705931
C	12	395.4	31.0	590	9	AU023153	AU023153	AU023153
C	13	393	30.8	450	10	BB946858	BB946858	UI-M-BH3-
C	14	392.2	30.7	419	10	BB704449	BB704449	BB704449
C	15	367.4	28.7	521	10	BB703859	BB703859	BB703859
C	16	367	28.8	427	10	BB706957	BB706957	BB706957
C	17	364.2	28.5	434	10	BB699732	BB699732	BB699732
C	18	347.2	27.2	391	10	BB704146	BB704146	BB704146
C	19	333.8	26.1	400	10	BB700620	BB700620	BB700620
C	20	319	25.0	905	28	B2114495	B2114495	CH230-421
C	21	315.2	24.9	441	10	BB704848	BB704848	BB704848
C	22	282.2	22.1	301	13	EX637052	EX637052	EX637052
C	23	264.6	20.7	315	10	BB248342	BB248342	BB248342
C	24	227.4	17.8	642	13	BY719704	BY719704	BY719704
C	25	224.8	17.6	654	28	A2225053	A2225053	R2C1-23-6
C	26	211.6	16.6	251	9	AV359580	AV359580	AV359580
C	27	208.6	16.3	251	9	AV358045	AV358045	AV358045
C	28	203.2	15.9	249	9	AV357766	AV357766	AV357766
C	29	203	15.9	290	10	BB063113	BB063113	BB063113
C	30	201.8	15.8	539	9	AL904756	AL904756	AL904756
C	31	200.8	15.7	909	13	EX755136	EX755136	EX755136
C	32	200	15.7	286	10	BB060381	BB060381	BB060381
C	33	199.8	15.6	246	10	BB060134	BB060134	BB060134
C	34	198	15.5	539	9	AL904892	AL904892	AL904892
C	35	197.8	15.5	642	12	BM651479	BM651479	BM651479
C	36	197	15.4	643	12	BU098465	BU098465	BU098465
C	37	196.8	15.4	526	9	AL904839	AL904839	AL904839
C	38	196.8	15.4	534	9	AL904820	AL904820	AL904820
C	39	196.8	15.4	559	9	AL904774	AL904774	AL904774
C	40	196.8	15.4	575	9	AL904855	AL904855	AL904855
C	41	196.4	15.4	575	9	AL904863	AL904863	AL904863
C	42	196.4	15.4	578	9	AL904817	AL904817	AL904817
C	43	196.4	15.4	579	9	AL904815	AL904815	AL904815
C	44	196.4	15.4	579	9	AL904816	AL904816	AL904816
C	45	196.4	15.4	604	9	AL904765	AL904765	AL904765

ALIGNMENTS

RESULT 1
LOCUS BB641267 675 bp mRNA linear EST 26-OCT-2001
DEFINITION BB641267 RIKEN full-length enriched, 10 days neonate cortex Mus
musculus cDNA clone AB30014H23 5', mRNA sequence.

ACCESSION BB641267
KEYWORDS BB641267.1 GI:16476392

ORGANISM Mus musculus (house mouse)

SOURCE Mus musculus

REFERENCE 1 (bases 1 to 675)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Arakawa,T., Carrinchi,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,C.,
Komoto,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ono,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

TITLE

Unpublished (2001)
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, S., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsubara, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Suganara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences *Mamm. Genome* 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

source

Location/Qualifiers

```
1. 666
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A830014H23"
/tissue_type="cortex"
/dev_stage="10 days neonate"
/lab_host="DH108"
/clone_1ib="RIKEN full-length enriched, 10 days neonate cortex"
```

/note="Site 1: Salt; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLC 1."

ORIGIN

Query Match 50.3%; Score 642.4; DB 10; Length 666;

Best Local Similarity 99.6%; Pred. No. 6.2e-87;

Matches 665; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
QY 588 GAGAGGCGATGTTCAAGCTGCAAGGAGGCGCGGTGGAGACAGACCAACCGAGAGA 647
Db 1 GAGAGGCGATGTTCAAGCTGCAAGGAGGCGCGGTGGAGACAGACCAACCGAGAGA 60
QY 648 CCGGAAACAGTGTGGCGCGGATGAGTCTGAGCTGGAGAGCGAGGCCATGTCCTGCGCC 707
Db 61 CCGGAAACAGTGTGGCGCGGATGAGTCTGAGCTGGAGAGCGAGGCCATGTCCTGCGCC 120
QY 708 AGAGATGGCTCAGAGACCCCGGTGATTCGATGCGCTCGAGACCAAGCCCTCCCGCAAG 767
Db 121 AGAGATGGCTCAGAGACCCCGGTGATTCGATGCGCTCGAGACCAAGCCCTCCCGCAAG 180
```

```
QY 768 CACGAGCAGACAGACAGAGCGCCTGCTGCTTCCAGTTCCTAGAGACAGAGAGTACGGCTACTA 827
Db 181 CACGAGCAGACAGACAGAGCGCCTGCTGCTTCCAGTTCCTAGAGACAGAGAGTACGGCTACTA 240
QY 828 TCAGTGCAGAGACCTGCAAAATCCGGTGGAGAGCGCCCTTATGTGTGTGTGACAGGGCAC 887
Db 241 TCAGTGCAGAGACCTGCAAAATCCGGTGGAGAGCGCCCTTATGTGTGTGTGACAGGGCAC 300
QY 888 CAGTAAGGTGTTACTTCAAAAGTTCGCCAGAGTGTGAGAAATCTTACACCCCTTACA 947
Db 301 CAGTAAGGTG-TACTTCAAAAGTTCGCCAGAGTGTGAGAAATCTTACACCCCTTACA 359
QY 948 GAGTGGAGACATCACTGTCGCAAAAGTTGTAAAGACTAGATGTGCTGCCAGTCAAGT 1007
Db 360 GAGTGGAGACATCACTGTCGCAAAAGTTGTAAAGACTAGATGTGCTGCCAGTCAAGT 419
QY 1008 TTGGCAGAGTGGAGACCTTAACGCCCGCATGGGAGAGCTGTGTGGAGATGCAAGAGCA 1067
Db 420 TTGGCAGAGTGGAGACCTTAACGCCCGCATGGGAGAGCTGTGTGGAGATGCAAGAGCA 479
QY 1068 AACGCTGTCTCTGAGAGACACCTTCAAGCTTCAATATCATTTAGTGAAGTCAAGAA 1127
Db 480 AACGCTGTCTCTGAGAGACACCTTCAAGCTTCAATATCATTTAGTGAAGTCAAGAA 539
QY 1128 CGTTTCGCTGATGAGGGGCTTAATGAAATGACAAAGTGAAGTCTTCCCTCTTCACTCT 1187
Db 540 CGTTTCGCTGATGAGGGGCTTAATGAAATGACAAAGTGAAGTCTTCCCTCTTCACTCT 599
QY 1188 TCCCTTCCAAATCTTCATAGACAGACAGAGTGTACTGTATTAAGCCTGTGAATAAA 1247
Db 600 TCCCTTCCAAATCTTCATAGACAGACAGAGTGTACTGTATTAAGCCTGTGAATAAA 658
QY 1248 GGTATTCG 1255
Db 659 GGTATTCG 666
```

RESULT 3

BC071693/c

LOCUS BC071693 615 bp mRNA linear EST 18-DEC-2003

DEFINITION H3102B03-3 NIA Mouse 15K cDNA clone Set Mus musculus cDNA clone

accession H3102B03 3', mRNA sequence.

VERSION BC071693

KEYWORDS BC071693.2 GI:40072037

EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 615)

Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grabovac, M.J., Parrano, S., Sano, Y., Piao, Y., Nagata, R., Doi, H., Wood, W.H., III, Becker, K.G. and Ko, M.S.H.

Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray

Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

TITLE

JOURNAL

MEDLINE

PUBMED

10922068

On Jan 26, 2001 this sequence version replaced gi:12554262.

Other_2STs: H3102B03-5

Contact: George J. Kargul

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igun.gsc.nia.nih.gov

This clone set has been freely distributed to the community. Please

visit <http://igun.gsc.nia.nih.gov/cdna/15k.html> for details.

Plate: H3102 row: B column: 03

Seq primer: -21M13 Forward

High quality sequence stop: 615

FEATURES

source

1. 615

Location/Qualifiers

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="nibest:H3102B03-3"
/db_xref="taxon:10090"
/clone="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI. This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extreme embryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extreme embryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

```

ORIGIN

```

Query Match      46.3%; Score 591.4; DB 10; Length 615;
Best Local Similarity 99.5%; Pred. No. 2.8e-79;
Matches 614; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 643 GAGGACCGGACAGTGTGGCGCGCATGTCAGTGTGAGCGAGGAGGAGCCATGTCCT 702
DB 615 GAGGACCGGACAGTGTGGCGCGCATGTCAGTGTGAGCGAGGAGGAGCCATGTCCT 556
QY 703 GCGCGAGAGTGGCTCAGGACCCCGGTGATTTGGATGTCCTCGAGACGAGGCTCCCG 762
DB 555 GCGCGAGAGTGGCTCAGGACCCCGGTGATTTGGATGTCCTCGAGACGAGGCTCCCG 496
QY 763 CAAAGGACGAGGACGAGGAGGAGGCGCTGCTCCAGTCTTAGAGCAGAGTAAAGGCG 822
DB 495 CAAAGGACGAGGACGAGGAGGAGGCGCTGCTCCAGTCTTAGAGCAGAGTAAAGGCG 436
QY 823 TACTATCACTGCAAGAGTGTCAAAATCCGATGAGAGCGGCTATGTGTGTGTGAG 882
DB 435 TACTATCACTGCAAGAGTGTCAAAATCCGATGAGAGCGGCTATGTGTGTGTGAG 376
QY 883 GGCACACGTAAGTGTACTTCAACAGTTCGCGGATGTGTGAGAAATCTTCAACCC 942
DB 375 GGCACACGTAAGTGT-TACTTCAACAGTTCGCGGATGTGTGAGAAATCTTCAACCC 317
QY 943 TTACAGAGTGAAGCAGATCACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCCAGT 1002
DB 316 TTACAGAGTGAAGCAGATCACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCCAGT 257
QY 1003 CAGATTTGCCCGACGTGAGACCCCTTAAAGCCGCCATGCGCAAGCTTGTGTGGAATGCA 1062
DB 256 CAGATTTGCCCGACGTGAGACCCCTTAAAGCCGCCATGCGCAAGCTTGTGTGGAATGCA 197
QY 1063 GAGCAAAACCCGTCTCTGCGACAGACCTTCAAGTTCAATATCATCTTATAGTGAATC 1122
DB 196 GAGCAAAACCCGTCTCTGCGACAGACCTTCAAGTTCAATATCATCTTATAGTGAATC 137
QY 1123 GAAAAAGTTTCTGCTAGATGGGCTTAAAGAAATGAGCAAGTGAAGTTTCTCCCTCTTCA 1182
DB 136 GAAAAAGTTTCTGCTAGATGGGCTTAAAGAAATGAGCAAGTGAAGTTTCTCCCTCTTCA 77
QY 1183 CCTCTTCCCTTCAAAATCTTCAATGACAGACAGTGTACTGATTAAGACCTGTGAA 1242

```

```

DB 76 CCTCTTCCCTTCAAAATCTTCAATGACAGACAGTGTACTGATTAAGACCTGTGAA 18
QY 1243 TAAAGTATGCAAC 1259
DB 17 TAAAGTATGCAAC 1

```

```

RESULT 4
LOCUS BC084538 595 bp mRNA linear EST 18-DEC-2003
DEFINITION H3102B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION BC084538
VERSION BC084538.2 GI:40072038
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

```

REFERENCE Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
1 (bases 1 to 595)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
Grabovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagareja,R., Doi,H.,
Wood,W.H., III, Becker,K.G., and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and
JOURNAL embryo using a 15,000 mouse developmental cDNA microarray
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
PUBMED 20381348
COMMENT On Jan 26, 2001 this sequence version replaced gi:12567102.
Other ESTs: H3102B03-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cnaa@leuun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
Plate: H3102 row: B column: 03
Seq primer: -21M3 Reverse
High quality sequence stop: 595
POLYA=No.

```

FEATURES

```

source
1..595
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="nibest:H3102B03-5"
/db_xref="taxon:10090"
/clone="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI. This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extreme embryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extreme embryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on

```

the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

Query Match 45.6%; Score 582.4; DB 12; Length 595;
Best Local Similarity 99.7%; Pred. No. 6.4e-78;
Matches 594; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 511 CCGGCAATCTCGGGAGACCCGGGAAACCGGAGCCGAGAGAGTGGCCCGGAGAAAGCGTGC 570
DB 1 CCGGCAATCTCTCGGGAGACCCGGGAAACCGGAGCCGAGAGAGTGGCCCGGAGAAAGCGTGC 60
QY 571 CCCGAGCCGAGAGACCGAGAGAGGAGTGTTCAGGCTGACAGGACCGGAGTGGAGAG 630
DB 61 CCCGAGCCGAGAGACCGAGAGAGGAGTGTTCAGGCTGACAGGAGCGCGGTGGAGAG 120
QY 631 CAGGCAATCTCGGGAGACCGGAAACAGTGTGGCGGAGTGTGAGTGTGAGCTGAGCTGAGAG 690
DB 121 CAGGCAATCTCTCGGGAGACCGGAAACAGTGTGGCGGAGTGTGAGTGTGAGCTGAGCTGAGAG 180
QY 691 GAGGCAATCTCTCGGGAGAGATGGCTCAGAGACCCCGGTATTCGATGCTCCCTCGAGAC 750
DB 181 GAGGCAATCTCTCGGGAGAGATGGCTCAGAGACCCCGGTATTCGATGCTCCCTCGAGAC 240
QY 751 CAGGCTCTCCCGGAAACCGAGAGACGAGAGAGAGGCTGCGTTTCCAGTCTTAGAG 810
DB 241 CAGGCTCTCCCGGAAACCGAGAGACGAGAGAGAGGCTGCGTTTCCAGTCTTAGAG 300
QY 811 CAGAGATGAGGCTATCTATCTACCTGAGAGACTGCAAAATCCGTGGAGAGCGCTATGTG 870
DB 301 CAGAGATGAGGCTATCTATCTACCTGAGAGACTGCAAAATCCGTGGAGAGCGCTATGTG 360
QY 871 TGGGTGTGAGAGGAGCAGCAGTAAAGTGTACTTCAAAACAGTGTGCGGAGTGTGAGAA 930
DB 361 TGGGTGTGAGAGGAGCAGCAGTAAAGTGTACTTCAAAACAGTGTGCGGAGTGTGAGAA 419
QY 931 ATCTCAACCCCTTACAGAGTGAAGACATCACTGTGCAAAAGTGTAAAGACTAGATG 990
DB 420 ATCTCAACCCCTTACAGAGTGAAGACATCACTGTGCAAAAGTGTAAAGACTAGATG 479
QY 991 TGCTTGGCCAGTCAAGATTTGGCCACGAGACCTTAAACGCCCTCAGGAGAGACTTGG 1050
DB 480 TGCTTGGCCAGTCAAGATTTGGCCACGAGACCTTAAACGCCCTCAGGAGAGACTTGG 539
QY 1051 TGGGAGATGAGAGAGCAAAAGCCTGTCTCTGAGAGAGAGAGCTTCAAAATACA 1106
DB 540 TGGGAGATGAGAGAGCAAAAGCCTGTCTCTGAGAGAGAGAGCTTCAAAATACA 595

RESULT 5 594 bp mRNA linear EST 04-DEC-2000
BF471866/c
LOCUS
DEFINITION
UT-M-BH3-awt-b-08-0-UT.1 NIH BMAP M S4 Mus musculus cDNA clone
UT-M-BH3-awt-b-08-0-UT 5', mRNA sequence.

ACCESSION
BF471866
VERSION
BF471866.1 GI:11541049
KEYWORDS
EST.
MUS musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
TITLE
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Mammalian Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890

Email: MEST@mail.nih.gov
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.
Location/Qualifiers
1. 594
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UT-M-BH3-awt-b-08-0-UT"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP M S4"
/note="Vector: pT7D-Fac (Pharmacia) with a modified
polylinker Site 1: Not I; Site 2: Eco RI; The
NIH BMAP M S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH BMAP M S4,
NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,
NIH BMAP M S2, NIH BMAP M S1. The subtracted library
(NIH BMAP M S4) was constructed as follows: PCR amplified
cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2 and
NIH BMAP M S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH BMAP M S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)"

ORIGIN

Query Match 44.7%; Score 570.4; DB 10; Length 594;
Best Local Similarity 99.5%; Pred. No. 4.1e-76;
Matches 593; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 555 CGCAGAGAAACGGTCCCGCAGCGAGAGAGAGGCGATGTTGAGTGCAGGCA 614
DB 594 CGCAGAGAAACGGTCCCGCAGCGAGAGAGAGGCGATGTTGAGTGCAGGCA 535
QY 615 GCGCGGTTGGAGAGACGACCAACCGAGAGACCGGAAACAGTGTGGCGGAGTGCAGTGC 674
DB 534 GCGCGGTTGGAGAGAGACGACCAACCGAGAGACCGGAAACAGTGTGGCGGAGTGCAGTGC 475
QY 675 TGAGCTTGGAGAGAGAGAGCAGTGTGCGCGAGAGATGCTCAGAGACCGCGGTGATTC 734
DB 474 TGAGCTTGGAGAGAGAGAGCAGTGTGCGCGAGAGATGCTCAGAGACCGCGGTGATTC 415
QY 735 GGAATGCCCTTGAAGACAGAGCTCCCGCAAAAGCAGAGAGAGAGAGCGCTCGC 794
DB 414 GGAATGCCCTTGAAGACAGAGCTCCCGCAAAAGCAGAGAGAGAGAGCGCTCGC 355
QY 795 TTTCAGTCTTGAAGAGAGAGTACGCTACTATCTGCAAGAGCTGCAAAATCCGGTG 854
DB 354 TTTCAGTCTTGAAGAGAGAGTACGCTACTATCTGCAAGAGCTGCAAAATCCGGTG 295
QY 855 GGAAGAGCGCTATGTGTGTGTGCAAGGAGCAGTAAAGTGTACTTCAAAAGTCT 914
DB 294 GGAAGAGCGCTATGTGTGTGTGCAAGGAGCAGTAAAGTGTACTTCAAAAGTCT 236

QY	915	GCCGAGTGTGTAGAAATCTCTCAACCCCTTCACAGAGTGGAGACATCACTGTCAAAGTT	974
Db	235	GCCGAGTGTGTAGAAATCTTCAACCCCTTCACAGAGTGGAGACATCACTGTCAAAGTT	176
QY	975	GTAAGAAACTAGATGTGCTGCTGCCACATTCCTGCACGTTGACCCCTTAAAGCCCCC	1034
Db	175	GTAAGAAACTAGATGTGCTGCTGCCACATTCCTGCACGTTGACCCCTTAAAGCCCCC	116
QY	1035	ATGCGCAAGACTTGTGTGGAGATGCAGACAAACGCTGTCTGTGGACAGCACTTCA	1094
Db	115	ATGCGCAAGACTTGTGTGGAGATGCAGACAAACGCTGTCTGTGGACAGCACTTCA	56
QY	1095	GCTTCAATACATCATTTAGTGGAGGTGAGAAAGTTCTGCTAGATGGGGGCTAAT	1150
Db	55	GCTTCAATACATCATTTAGTGGAGGTGAGAAAC-TTTGTGCTAGATGGGGGCTAAT	1

RESULT 6	
BB704019	
LOCUS	525 bp, mRNA linear, EST 11-OCT-2001
DEFINITION	BB704019 RIKEN full-length cDNA clone, in vitro fertilized eggs Mus musculus cDNA clone 7420459B08 3', mRNA sequence.
ACCESSION	BB704019
VERSION	BB704019.1 GI:16052854
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1. (Baes 1 to 525)
AUTHORS	Akimura,T., Aakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,

TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.

JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki

FEATURES

Source

```

1..525
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10930"
/clone="7420459B08"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="Riken full-length enriched, in vitro
fertilized eggs"
/name="Site_1; Sali; Site_2; BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5']
GAGGAGAGAGAGATTCGACCAAGGCTCTTTTCTTTTCTTTTCTTNN 3'}. cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5']
GAGGAGAGAGATTCGAGATTCTTAATTAATTAATCCCCCCCCCCCC 3}. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified plasmid pSP65(+) after bulk excision from Lambda
phage 1. Cloning sites, 5' end:Sali; 3' end: BamHI"

```

ORIGIN

Query Match	38.9%	Score 497.2;	DB 10;	Length 525;
Best Local Similarity	99.0%;	Pred. No. 4.1e-65;		
Matches 521;	Conservative 0;	Mismatches 3;	Indels 2;	Gaps 2;

[illegible]

RESULT 7
AI854700/c

LOCUS	AI854700	499 bp	MRNA	linear	EST 15-JUL-1995
DEFINITION	UI-M-BH0-akc-d-12-0-UI-s1 NIH BMP4	MI	Mus musculus	cDNA clone	
ACCESSION	UI-M-BH0-akc-d-12-0-UI 3'			mRNA sequence.	
VERSION	AI854700				
KEYWORDS	AI854700.1	GI:5498606			
SOURCE	EST.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Mus musculus				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 499)				
COMMENT	Bonardo, M.F., Lennon, G. and Soares, M.B.				
	Normalisation and subtraction: two approaches to facilitate gene discovery				
	Genome Res. 6 (9), 791-806 (1996)				
JOURNAL	97044477				
MEDLINE	8889548				
COMMENT	Contact: Chin, H				

Email: m85t@mail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized basal ganglia library cDNA library Preparation: M.B. Soares/ABD Clonte distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
SOLVA=Yes.

```

FEATURES
source

Location/Qualifiers
1. 499
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH0-alc-d-12-0-UI"
/dev_stage="27-32 days"
/lab_host="PH108 (Life Technologies)"
/clone_lib="NH_BMAP_M_S1"
/notes="Vector: pTRIP-Pac (Pharmacia) with a modified
polylinker: Site 1: Not I; Site 2: Eco RI; The
NH_BMAP_M_S1 library is a subcloned library derived from
a mixture of normalized libraries from ten regions of the
mouse brain (cerebellum, brain stems, olfactory bulbs,
hippocampus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus). The driver used for
subcloning consisted of a pool of 20,000 cDNA clones
obtained from non-normalized and normalized libraries of
these ten regions of the mouse brain.
TNG_TISSUE=basal ganglia
TNG_LIB=NH_BMAP_M_S1
TAG_SEQ=GTATC"

```

ORIGIN	Query Match	Best Local Similarity	Matches	Score	DB %	Length	Mismatches	Indels	Gaps
	37.4%	99.2%	497	473.8	1.3e-61	499	0	2	2
Qy	772	GAGAGAGACAGAGAGCGCCGCGCTTCCAGTCTTAGAGACAGAAATACGCTACTATAC	499	GAGCAGACACAGAGAGCGCCTCGCTTCCAGTCTTAGAGACAGAAATACGCTACTATAC	832				
Db	499	GAGCAGACACAGAGAGCGCCTCGCTTCCAGTCTTAGAGACAGAAATACGCTACTATAC	499	GAGCAGACACAGAGAGCGCCTCGCTTCCAGTCTTAGAGACAGAAATACGCTACTATAC	440				
Qy	832	TGCAGAGACTGCAGAAATCCGGTGGAGAGCGCCTATGTGTGTGTGTGCAGGACCAAGT	439	TGCAGAGACTGCAGAAATCCGGTGGAGAGCGCCTATGTGTGTGTGTGCAGGACCAAGT	380				

QY	892	AAGGTGTACTTCAAAAGATTCTCCGAGTGTGAGAAATCCTACAAACCTTACAGAGT	951
Dp	379	AAGGTG-TACTTCAAAAGATTCTCCGAGTGTGAGAAATCCTACAAACCTTACAGAGT	321
QY	952	GGAGGACATCACTGTCAAAGTTGTAAAGAAGTCTAGATGTGCTGCCAGTCAGATTTCG	1011
Dp	320	GGAGGACATCACTGTCAAAGTTGTAAAGAAGTCTAGATGTGCTGCCAGTCAGATTTCG	261
QY	1012	CCAGGTGAGCCCTTAAGCCGCCCATTCGGCAAGATTGTGTGGAGATGCAAGACMAACG	1071
Dp	260	CCAGGTGAGCCCTTAAGCCGCCCATTCGGCAAGATTGTGTGGAGATGCAAGACMAACG	201
QY	1072	CCGTGCTCGGAGACGACCTTCGACTTCAATACATCATTTAGAGAGTGGAAAAGTT	1131
Dp	200	CCGTGCTCGGAGACGACCTTCGACTTCAATACATCATTTAGAGAGTGGAAAAGTT	141
QY	1132	TCTGTGATGTGGGGCTAATGGAATGGAAGAAGTCTTCCCTCTTCACTCTTCCC	1191
Dp	140	TCTGTGATGTGGGGCTAATGGAATGGAAGAAGTAGCTTTCTCCCTCTTCACTCTTCCC	81
QY	1192	TTTTCGAAATTCCTCATGACAGACAGTGTACTTGGATATTAAGCCTGTGAATTAAGGTA	1251
Dp	80	TTTTCGAAATTCCTCATGACAGACAGTGT-TACTTGGATATTAAGCCTGTGAATTAAGGTA	22
QY	1252	TTTTCGAAATTCGAAAAAAAAAAAAA	1272
Dp	21	TTTTCGAAATTCGAAAAAAAAAAAAA	1

RESULT 8					
BB703259					
LOCUS	BB703259	491 bp	mRNA	linear	EST 11-OCT-2001
DEFINITION	BB703259	RIKEN full-length enriched	in vitro fertilized eggs	Mus	
	musculus	cDNA clone 7420449J15	3'	mRNA	sequence.
ACCESSION	BB703259				
VERSION	BB703259.1	GI:16052094			

KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Mus. 1 (bases 1 to 491)
AUTHORS	Akimura, T., Aikawa, T., Carinci, P., Furuno, M., Hanagaki, T., Hayata, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imutani, K., Ishii, Y., Ito, M., Kawai, J., Koizumi, Y., Komuro, H., Kouda, R., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, M., Ohashi, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, I., Sato, K., Shibata, K., Shitahara, A., Shitahara, A., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, I., Tomaru, A., Toyota, I., Watahiki, A., Yasunishi, A., Yumetani, M., and Hayashizaki, Y.
TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shohiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 226-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komu, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
wagdi, K., Fujisawa, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E.,
Watabiki, M., Yoneda, Y., Ishikawa, T., Osawa, K., Tanaka, T.,
Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format

/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: Sali; Site 2: BamHI; CDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATCTCGAGTTATTAATTAATCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified Bluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

ORIGIN

Query Match 32.4%; Score 413.8; DB 10; Length 491;
Best Local Similarity 93.0%; Pred. No. 1,46-52;
Matches 455; Conservative 0; Mismatches 32; Indels 2; Gaps 2;

QY 768 CACGAGAGAGACAGAGAGCGCTGCTTCCAGTTCTTGAAGAGAGAGAGCTACTA 827
Db 1 CACGAGAGAGACAGAGAGCGCTGCTTCCAGTTCTTGAAGAGAGAGAGCTACTA 60
QY 828 TCACTGCAAGAGACTGCAGAAATCCGATGAGAGAGCGCTATGTGTGTGTGAGAGGAC 887
Db 61 TCA-AGCAAGAGACTGCAGAAATCCGATGAGAGAGCGCTATGTGTGTGTGAGAGGAC 119
QY 888 CAGTAAGGTGTACTTCAAAACAGTTTCGCGAGGTGTGAGAAATCCTTAACCTTACA 947
Db 120 CAGTAAGGTGTACTTCAAAACAGTTTCGCGAGGTGTGAGAAATCCTTAACCTTACA 179
QY 948 GAGTGAAGAGATCAGCTGTCAAAAGTTGTAAGAAATGATGTGCTGCCAGTCAAT 1007
Db 180 GAGTGAAGAGATCAGCTGTCAAAAGTTGTAAGAAATGATGTGCTGCCAGTCAAT 239
QY 1008 TTGCGCACGTGAGACCTTAAAGCGCCCATGCGCAAGCTTGTGTGAGATGCAAGACA 1067
Db 240 TTGCGCACGTGAGACCTTAAAGCGCCCATGCGCAAGCTTGTGTGAGATGCAAGACA 299
QY 1068 AACGCTGTCTGTGAGAGAGAGCTTGAAGTCAATCATCTTATGAGAGTCAAAA 1127
Db 300 AACGCTGTCTGTGAGAGAGAGCTTGAAGTCAATCATCTTATGAGAGTCAAAA 359
QY 1128 CGTTTGTCTGATGAGAGAGCTTGAAGTCAATCATCTTATGAGAGTCAAAA 1187
Db 360 CGTTTGTCTGATGAGAGAGCTTGAAGTCAATCATCTTATGAGAGTCAAAA 419
QY 1188 TCCCTTCCAAATCTTCATGAGAGAGAGTGTATGATTAAGAGCTGTAATAAA 1247
Db 420 TCCCTTCCAAATCTTCATGAGAGAGAGTGTATGATTAAGAGCTGTAATAAA 478
QY 1248 GGTATTGCA 1256
Db 479 GGTATTGTA 487

RESULT 10
A0044294/c 436 bp mRNA linear EST 09-DEC-1998
LOCUS A0044294 Mouse sixteen-cell-embryo cDNA Mus musculus cDNA clone
DEFINITION J0917609 3', mRNA sequence.
ACCESSION A0044294
VERSION A0044294.1 GI:3979844
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. '436
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="J0917609"
/dev_stage="sixteen-cell-embryo"
/clone_lib="Mouse sixteen-cell-embryo cDNA"

ORIGIN

Query Match 31.8%; Score 405.8; DB 9; Length 436;
Best Local Similarity 97.0%; Pred. No. 2,2e-51;
Matches 424; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 823 TACTATCACTGCAAGAGACTGCAAAATCCGTGGAGAGAGCGCTATGTGTGTGTGAG 882
Db 436 TACTATCACTGCAAGAGACTGCAAAATCCGTGGAGAGAGCGCTATGTGTGTGTGAG 377
QY 883 GGCACAGTAAGGTGTCTTCAACAGTTCTGCCAGTGTGTGAGAAATCCTTAACCC 942
Db 376 GGCACAGTAAGGTGTCTTCAACAGTTCTGCCAGTGTGTGAGAAATCCTTAACCC 317
QY 943 TTACAGAGTGAAGACATCACCTGTCAAAAGTTGTAAGAAATGATGTGCTGCCAGT 1002
Db 316 TTACAGAGTGAAGACATCACCTGTCAAAAGTTGTAAGAAATGATGTGCTGCCAGT 257
QY 1003 CAGATTTGGCCACGTGAGACCTTAAAGCGCCCATGCGCAAGACTGTGTGAGATGCA 1062
Db 256 CAGATTTGGCCACGTGAGACCTTAAAGCGCCCATGCGCAAGACTGTGTGAGATGCA 197
QY 1063 GGACAAAGCGCTGTCTGCGACAGACCTTCAATATCATCTTATGAGAGTCA 1122
Db 196 GGACAAAGCGCTGTCTGCGACAGACCTTCAATATCATCTTATGAGAGTCA 137
QY 1123 GAAAACGTTTCTGTAGTGGGGCTAATGAAATGAGCAAGAGCTTCTCCCTCTTCA 1182
Db 136 GAAAACGTTTCTGTAGTGGGGCTAATGAAATGAGCAAGAGCTTCTCCCTCTTCA 77
QY 1183 CCTTTCCTTCCAAATCTTCAATGAGAGAGAGTGTATGATTAAGAGCTGTGAA 1242
Db 76 CCTTTCCTTCCAAATCTTCAATGAGAGAGAGTGTATGATTAAGAGCTGTGAA 18
QY 1243 TAAAAGTATTGCAAC 1259
Db 17 TAAAAGTATTGCAAC 1

RESULT 11
BB705931 425 bp mRNA linear EST 11-OCT-2001
LOCUS BB705931 RIKEN full-length enriched, in vitro fertilized eggs Mus
DEFINITION BB705931 Musculus cDNA clone 7420484F16 3', mRNA sequence.
ACCESSION BB705931
VERSION BB705931.1 GI:16054766
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RESULT 14			
LOCUS	BE704449	419 bp	mRNA linear EST 11-OCT-2001
DEFINITION	BE704449 RIKEN full-length enriched, in vitro fertilized eggs Mus musculus cDNA clone 7420464A04 3', mRNA sequence.		
ACCESSION	BE704449		
VERSION	BE704449.1		
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 419)		
AUTHORS	Altamura, T., Arikawa, T., Carninci, P., Furuno, M., Hasegaki, T., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Iihii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ozazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watanishi, A., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.		
TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Yoshinori Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenro-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-researc.riken.go.jp, url:http://genome.gsc.riken.go.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10) 1617-1630 (2000) Wagci, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, T., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11) 1757-1771 (2000) Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. e mouse tissues. Location/Qualifiers 1. 419 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="7420464A04" /sex="female" /tissue_type="in vitro fertilized eggs" /dev_stage="egg" /lab_host="DH10B" /clone_lib="RIKEN full-length enriched, in vitro fertilized eggs" /note="Site 1: Salt; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken		

	Query Match	30.7%	Score 392.2;	DE 10;	Length 419;					
	Best Local Similarity	98.8%;	Pred. No. 2.5e-49;							
	Matches 416;	Conservative 0;	Mismatches 3;	Indels 2;	Gaps 2;					
ORIGIN										
QY	834 CAAGACTGCAAAATCCGGTGGAGAGCGCCTATGTGTGTGTGTGCAGGGCACACAGTAA	893								
Dd	1 CAAAGACTGCAAAATCCGGTGGAGAGCGCCTATGTGTGTGTGTGCAGGGCACACAGTAA	60								
QY	894 GGTGTTAATTCAAACAGTTCTTGCCGAGTGTGTGAANAATCTTACAACCCTTAACAGATGG	953								
Dd	61 GGT-TTACTTCAAAACAGTTCTTGCCGAGTGTGTGAANAATCTTACAACCCTTAACAGATGG	119								
QY	954 AGGACATCACCTGTCCAAGAATGTTAAAGAACTAATATGTCCTGCCAGTCAGATTTCCGC	1013								
Dd	120 AGGACATCACCTGTCCAAGAATGTTAAAGAACTAATATGTCCTGCCAGTCAGATTTCCGC	179								
QY	1014 ACGTGGACCTTAAACGCCCCCATATGGGCACAGACTTGTGTGGAGATGCAAGAACAAAGCCC	1073								
Dd	180 ACGTGGACCTTAAACGCCCCCATATGGGCACAGACTTGTGTGGAGATGCAAGAACAAAGCCC	239								
QY	1074 TGTCTCTGGGACAGACACCTTCAGCTTCAAATACATCATTTAGTGTGAGATGCGAAAAAGTTTC	1133								
Dd	240 TGTCTCTGGGACAGACACCTTCAGCTTCAAATACATCATTTAGTGTGAGATGCGAAAAAGTTTC	299								
QY	1134 TGCTTAGATGGGGCTTAATGGAATGSAACAAGTGAAGCTTTCTCCCCCTTTTACCTCTTCCCTT	1193								
Dd	300 TGCTTAGATGGGGCTTAATGGAATGSAACAAGTGAAGCTTTCTCCCCCTTTTACCTCTTCCCTT	359								
QY	1194 TCCTAAATTCCTTCATGACGACAGAGTGTACTGTGATATAAAGCCTGTGAATAAAGATATT	1253								
Dd	360 TCCTAAATTCCTTCATGACGACAGAGTGTACTGTGATATAAAGCCTGTGAATAAAGATATT	418								
QY	1254 G 1254									
Dd	419 G 419									
RESULT 15										
BH703869										
LOCUS	BH703869	521 bp	mRNA	linear	EST 11-OCT-2001					
DEFINITION	Musculus full-length enriched, in vitro fertilized eggs Mus									
ACCESSION	musculus cDNA clone 7420457C21.3, mRNA sequence.									
VERSION	BH703869									
KEYWORDS	BH703869.1 GI:16052704									
SOURCE	EST.									
ORGANISM	Mus musculus (house mouse)									
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
AUTHORS	Akimura,T., Arakawa,T., Garinaci,P., Furuno,M., Hanagaki,T., Hayabara,N., Hiramoto,K., Hirooka,T., Hitozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouada,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido.T., Saito.R., Sakai.C., Sakai.K., Sakazume.N., Saeki.D., Sato.K., Shibata.K., Shinagawa,A., Shiraki.T., Sogabe.Y., Suzuki.H., Tagawa.A., Takahashi.F., Takaku-Akahira,S., Tanaka.T., Tomaru.A., Toyo.T., Wataniki,A., Yasunishi,A., Yamamatsu.M. and Hayashizaki,Y.									
TITLE	Murine Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.									

the oocyte-to-embryo transition
 Nat. Genet. 33 (2), 187-191 (2003)
 MEDLINE 22447938
 PUBMED 12539046
 REFERENCE 2 (bases 1 to 1260)
 AUTHORS Wu, X., Wang, P. and Matzuk, M.M.
 TITLE Direct Submission
 JOURNAL Submitted (04-Dec-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 FEATURES
 Source
 1..1260
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/SvEv"
 /db_xref="taxon:10090"
 /chromosome="5"
 /map="between DBuc48 and Txk"
 1..1260
 /gene="Zarl"
 1..27
 /gene="Zarl"
 28..1113
 /gene="Zarl"
 /note="Zarl; maternal factor; contains atypical C9 PHD motif at C-terminus"
 /codon_start=1
 /product="zygote arrest 1"
 /protein_id="AA024706.1"
 /db_xref="gi:27808690"
 /translation="MFPASTFPCPHYPQATKAGDMRFEGACBPAPSPFPGYRQ LMAEYDSHROALMALISMGRSVSSDAAYVNRDASVCSLSEVAGSOTPKGCS RASPDAGSGCOPRGAAGSPRSWOTVAFSIVTCRGSLSLEVAAGSOTPKGCS SPASGTRPREPRVARARKVAPRPSREGEYQAAAGMGMOQPPEDRNSVAAWSEF GSSEPCRAEMADPDPSDAPRQASQSTQDERKRFQFLQKGYHCKQCKLRM ESAYVWCQVGSKVYKFCVCKSYNPRVEDITQSCRTACAPVALRHVDKPR PHRDLCGRCKRSLSDSTSFXYII"
 1114..1260
 /gene="Zarl"

3' UTR
 ORIGIN

Query Match 97.5% Score 1245.4; DB 10; Length 1260;
 Best Local Similarity 99.8% Pred. No. 2; Be-248;
 Matches 1257; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 3 GGGGGGAGGGGAGGGGAGGACCATGTTCCCGGAGAGACGTTCCACCCCTGACCGCA 62
 3 GGGGGGAGGGGAGGGGAGGACCATGTTCCCGGAGAGACGTTCCACCCCTGACCGCA 62
 63 TCCTTATCCGAGGACCAACAAAGCGGAGATGCTGAGAGTTCCGAGCCAGGGGCTGCCG 122
 63 TCCTTATCCGAGGACCAACAAAGCGGAGATGCTGAGAGTTCCGAGCCAGGGGCTGCCG 122
 123 ACCGGGGGGGGCT 182
 123 ACCGGGGGGGGCT 182
 183 CAGCCACAGGGGAGGACGATCATGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
 183 CAGCCACAGGGGAGGACGATCATGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
 243 CAGCCGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 302
 243 CAGCCGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 302
 303 CAGCCGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 362
 303 CAGCCGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 362
 363 CTGTCAACCCCGTGGCCACCGCCGGGGAGATCCCGCCGATCTCTGGAGACCGCTAGC 422
 363 CTGTCAACCCCGTGGCCACCGCCGGGGAGATCCCGCCGATCTCTGGAGACCGCTAGC 422
 423 CCCGTTCTGTGCTGACCTTCTGTGGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482

Db 423 CCCGTTCTGTGCTGACCTTCTGTGGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 Qy 483 GCGAGACACCAAG 542
 Db 483 GCGAGACACCAAG 542
 Qy 543 GAGAGAGTGGCGCGAG 602
 Db 543 GAGAGAGTGGCGCGAG 602
 Qy 603 GGTGTGAGGAGGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
 Db 603 GGTGTGAGGAGGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
 Qy 663 GCGAGATGACGTCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
 Db 663 GCGAGATGACGTCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
 Qy 723 CCCCGGTGATTCGAGATGCCCTTCGAGACCAAGGCTCCCGCAAGCAGAGAGAGAG 782
 Db 723 CCCCGGTGATTCGAGATGCCCTTCGAGACCAAGGCTCCCGCAAGCAGAGAGAGAG 782
 Qy 783 GAGAGGCGTGGGCTTCCAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
 Db 783 GAGAGGCGTGGGCTTCCAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
 Qy 843 CAAATCCCGTGGAGAGAGGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902
 Db 843 CAAATCCCGTGGAGAGAGGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 901
 Qy 903 TCAACAGTTCGCGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 962
 Db 903 TCAACAGTTCGCGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 961
 Qy 963 CCTGTCAAGTGTCAAG 1022
 Db 963 CCTGTCAAGTGTCAAG 1021
 Qy 1023 CTAAACGCGCCCATGCGCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1082
 Db 1023 CTAAACGCGCCCATGCGCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1081
 Qy 1083 ACAGCAGCTTCAAGCTTCAATATCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1142
 Db 1083 ACAGCAGCTTCAAGCTTCAATATCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1141
 Qy 1143 GGGCTAATGGAATGAGCAAGTCTTCTCCCTTTCACCTCTTCCCTTCCAAATTC 1202
 Db 1143 GGGCTAATGGAATGAGCAAGTCTTCTCCCTTTCACCTCTTCCCTTCCAAATTC 1201
 Qy 1203 TTCAATGACAGACAGTGTACTTGAATTAAGCCTGTGAATTAAGGATTCGAATCA 1261
 Db 1203 TTCAATGACAGACAGTGTACTTGAATTAAGCCTGTGAATTAAGGATTCGAATCA 1260

RESULT 2
 AY283175 1280 bp mRNA linear ROD 21-Aug-2003
 LOCUS Rattus norvegicus zygote arrest 1 (Zarl) mRNA, complete cds.
 DEFINITION AY283175
 ACCESSION AY283175.1 GI:30908932
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus;
 1 (bases 1 to 1280)
 Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.
 Zygote arrest 1 (Zarl) is an evolutionarily conserved gene
 expressed in vertebrate ovaries
 JOURNAL Biol. Reprod. 69 (3), 861-867 (2003)

TITLE JOURNAL REFERENCE AUTHORS

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 147464)
 Biren, B., Nusbaum, C., Lander, E., Abouelella, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Erickson, J., Fano, S., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Ekstrom, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Haler, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanab, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Madditt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meidrim, J., Menus, L., Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, V., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (02-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 2, 2003 this sequence version replaced g121206317.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 Project Information
 Center project name: L17506
 Center clone name: 506_B_15
 Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 140214 bases at least Q40
 Consensus quality: 143454 bases at least Q30
 Consensus quality: 144691 bases at least Q20
 Insert size: 14400; agarose-fp
 Insert size: 145464; sum-of-coverage
 Quality coverage: 4.6 in Q20 bases; agarose-fp
 Quality coverage: 4.6 in Q20 bases; sum-of-coverage
 NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 1
 654 753: gap of 653 bp in length
 754 1934: contig of 1181 bp in length
 1935 2034: gap of 100 bp
 31965: contig of 29662 bp in length
 31997 32096: gap of 100 bp
 32097 32986: contig of 890 bp in length
 32987 33086: gap of 100 bp
 33087 33811: contig of 725 bp in length
 33812 33911: gap of 100 bp
 33912 35429: contig of 1518 bp in length
 35430 35529: gap of 100 bp

FEATURES
 source
 1..147464
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone_id="RP24-506B15"
 /clone_lib="RPCI-24 Male Mouse BAC"
 1..653
 /note="assembly_fragment"
 754..1934
 /note="assembly_fragment"
 2035..31996
 /note="assembly_fragment"
 32097..32986
 /note="assembly_fragment"
 33087..33811
 /note="assembly_fragment"
 33912..35429
 /note="assembly_fragment"
 35530..38157
 /note="assembly_fragment"
 38258..41538
 /note="assembly_fragment"
 41639..43241
 /note="assembly_fragment"
 43342..46930
 /note="assembly_fragment"
 47031..51340
 /note="assembly_fragment"
 51441..56135
 /note="assembly_fragment"
 56236..61691
 /note="assembly_fragment"
 61792..70053
 /note="assembly_fragment"
 70154..76663
 /note="assembly_fragment"
 76764..86117
 /note="assembly_fragment"
 86218..94174
 /note="assembly_fragment"
 94275..108523
 /note="assembly_fragment"
 108624..123796
 /note="assembly_fragment"
 123897..146961
 /note="assembly_fragment"
 147062..147464
 /note="assembly_fragment"

misc_feature
 /note="assembly_fragment"
 754..1934
 /note="assembly_fragment"
 2035..31996
 /note="assembly_fragment"
 32097..32986
 /note="assembly_fragment"
 33087..33811
 /note="assembly_fragment"
 33912..35429
 /note="assembly_fragment"
 35530..38157
 /note="assembly_fragment"
 38258..41538
 /note="assembly_fragment"
 41639..43241
 /note="assembly_fragment"
 43342..46930
 /note="assembly_fragment"
 47031..51340
 /note="assembly_fragment"
 51441..56135
 /note="assembly_fragment"
 56236..61691
 /note="assembly_fragment"
 61792..70053
 /note="assembly_fragment"
 70154..76663
 /note="assembly_fragment"
 76764..86117
 /note="assembly_fragment"
 86218..94174
 /note="assembly_fragment"
 94275..108523
 /note="assembly_fragment"
 108624..123796
 /note="assembly_fragment"

Db	131304	ACCCGCGCCCCCTCTCTCTCTCCCGCGATACAGACAGCTCATGCGCCGGGAGTACGTGCA	131366
OY	183	CAGGCAACCAACGGGAGCAAGCTATAGGCCCTGCTGTGTGGGAGATGGGTCCCCGGTGGTCAg	242
Db	131364	CAGGCACCAAGGGGCAACGTCAATGAGCCCTGTGCTGTGGAGATGGGTCTCCCGTGGTCAg	131423
OY	243	CAGCCTGTAGAGCTGTGCGGTGACAGATGAACCCGGCCGCGACAGCTCTGGTGCAGTGTCACT	302
Db	131424	CAGCCTGTAGAGCTGTGCGGTGAGAGTGAACCGCGCCGCGACAGCTCTGGTGCAGTGTCACT	131485
OY	303	CGGCGCCCGCAACGTTGAGCTTGTGAGCTTGTGAGGGTGTGCCGAGCCCGACGCCCGATGGGTTT	362
Db	131484	CGGCGCCCGCAACGTTGAGCTTGTGAGCTTGTGAGGGTGTGCCGAGCCCGACGCCCGATGGGTTT	131543
OY	363	CTGTCAACCCCGTGGGCGACGCGGCGCGCGGAGATATCCCGGAGATCTCTGACAGACCGTAg	422
Db	131544	CTGTCAACCCCGTGGGCGACGCGGCGCGCGGAGATATCCCGGAGATCTCTGACAGACCGTAg	131603
OY	423	CCCGTCTCGTCCGTGACCTTCTGTGGCCTCTCTCTCACTGAGAGTTGCGGAGGCGAG	482
Db	131604	CCCGTCTCGTCCGTGACCTTCTGTGGCCTCTCTCTCACTGAGAGTTGCGGAGGCGAG	131665
OY	483	GCAACACCCACGAAAGGAGAGAGGGAGCCCGGCGATCTTGTGGGAGACCCGGGAAACCGAGCC	542
Db	131664	GCAACACCCACGAAAGGAGAGAGGGAGCCCGGCGATCTTGTGGGAGACCCGGGAAACCGAGCC	131723
OY	543	GAGAGAGTGTGGCCGCGAGAAACCGGTCTCCCGACGCGCGAAGCGAGAGGCGCGATGTTC	602
Db	131724	GAGAGAGTGTGGCCGCGAGAAACCGGTCTCCCGACGCGCGAAGCGAGAGGCGCGATGTTC	131783
OY	603	GGCTGCAGGGCGAGCCGGGTGGGAGAGCAGCCACCCAGAGAGCCGGAAACAGTGTGCG	662
Db	131784	GGCTGCAGGGCGAGCCGGGTGGGAGAGCAGCCACCCAGAGAGCCGGAAACAGTGTGTGC	131843
OY	663	GGCGATCACTCTGAGCCTGTGGAGCGAGAGGACATGTCTCTGCGAGAGATGGGTCAGGA	722
Db	131844	GGCGATCACTCTGAGCCTGTGGAGCGAGAGGACATGTCTCTGCGAGAGATGGGTCAGGA	131903
OY	723	CCCCGATGATTGGATGCCCTTGAGAACCGAGGCTTCCCGCAAAACGAGACAGAGCAA	782
Db	131904	CCCCGATGATTGGATGCCCTTGAGAACCGAGGCTTCCCGCAAAACGAGACAGAGCAA	131963
OY	783	GGAGCGCTGCGCTTCCAGTT	803
Db	131964	GGAGCGCTGCGCTTCCAGTT	131984
RESULT 6			
AC107686		168073 bp	DNA linear HTG 19-NOV-2003
LOCUS			
DEFINITION		Mus musculus chromosome 5 clone RP24-228B12 map 5, WORKING DRAFT	
ACCESSION		AC107686	
VERSION		AC107686.3	
KEYWORDS		HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULFILL	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Buatolet, Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		1 (bases 1 to 168073)	
TITLE		1 (bases 1 to 168073)	
JOURNAL		Unpublished	
AUTHORS		2 (bases 1 to 168073)	
		Bitren, B., Lincon, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collins, A., Cook, A., Cooke, P., DeArlelano, K., Dewar, K., Diaz, J.S., Dodde, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Larocque, K., Lamazares, R.,	

TITLE	Direct Submission
JOURNAL	Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 168073)
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Beabhi, H. W., Bantz, N., Bhat, T.,

TITLE	Direct Submission
JOURNAL	Submitted (19-NOV-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Nov 19, 2003 this sequence version replaced gi:20258450.

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: <http://www-seq.wi.mit.edu>

----- Project Information -----

Center clone name: 228_B_12

Summary Statistics

Chemistry: Dye-terminator Big Dye; 100% c

```
Assembly program: Phrap; version 0.960731
```

Consensus quality: 167176 bases at least Q30

Consensus quality: 167318 bases at least Q20

Insert size: 167473; sum-of-contigs

Quality coverage: 12.4 in Q20 bases; agarose-tp

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

NOTE: This is a 'working draft' sequence. It currently consists of 7 sections. Cons between the sections

re represented as runs of N . The order of the pieces followed by the number of runs between the first

is believed to be correct as given, however the sizes

provided by the submittor.

by the finished sequence as soon as it is available and

the accession number will be preserved.

```

*      1      3391: contig of 3391 bp in length
*      3392      3491: gap of 100 bp
*      3492      11221: contig of 7730 bp in length
*      11222      11321: gap of 100 bp
*      11322      20171: contig of 8850 bp in length
*      20172      20271: gap of 100 bp
*      20272      45821: contig of 25550 bp in length
*      45822      45921: gap of 100 bp
*      45922      100777: contig of 54856 bp in length
*      100778      100877: gap of 100 bp
*      100878      143469: contig of 42592 bp in length
*      143470      143569: gap of 100 bp
*      143570      168073: contig of 24504 bp in length.
*
FEATURES
Source
1..168073
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP24-228B12"
/clone_id="RPCI-24 Male Mouse BAC"
1..3331
/note="assembly_fragment"
/clone_end:SP6
/vector_side:left"
misc_feature
3492..11221
/note="assembly_fragment"
11322..20171
/note="assembly_fragment"
20272..45821
/note="assembly_fragment"
45922..100777
/note="assembly_fragment"
100878..143469
/note="assembly_fragment"
143570..168073
/note="assembly_fragment"
clone_end:T7
/vector_side:right"
ORIGIN
Query Match      54.0%; Score 689.8; DB 2; Length 168073;
Best Local Similarity 95.0%; Pred. No. 6,7e+133;
Matches 761; Conservative 0; Mismatches 22; Indels 16; Gaps 4;
QY      3      GGGCGGGGAGGCGGGGAGCGACCCATGTTCCCGGCGAGACGTTCCACCCCTGCCGCA 62
Db      63451  GGGCGGGGAGGCGGGGAGCGACCCATGTTCCCGGCGAGACGTTCCACCCCTGCCGCA 63510
QY      63      TCCTTATCCGAGGCGCCAAAGCGGGGATGCTGAGGTGGAGGCGAGGCGGCTGCGG 122
Db      63511  TCCTTATCCGAGGCGCCAAAGCGGGGATGCTGAGGTGGAGGCGAGGCGGCTGCGG 63570
QY      123      ACCCGGGCCCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTCA 182
Db      63571  ACCCGGGCCCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTCA 63630
QY      183      CAGGCGACGAGGCGGCAAGCTAGCGGCTGCGGCGGATGGGTCCCGGCTGCGTCA 242
Db      63631  CAGGCGACGAGGCGGCAAGCTAGCGGCTGCGGCGGATGGGTCCCGGCTGCGTCA 63690
QY      243      CAGCGGTGAGCGTGCAGGTGAGTGAACCCGCGCGCGAGCGCTCGGTGAGTGTCACT 302
Db      63691  CAGCGGTGAGCGTGCAGGTGAGTGAACCCGCGCGCGAGCGCTCGGTGAGTGTCACT 63750
QY      303      CGGGCGCGGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 362
Db      63751  CGGGCGCGGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 63810
QY      363      CTGTCAACCCCGTGGCCACCGCGCGCGGAGATCCCGCGAGTCTGTCAGACGCTAGC 422
Db      63811  CTGTCAACCCCGTGGCCACCGCGCGCGGAGATCCCGCGAGTCTGTCAGACGCTAGC 63870

```

```

QY      423      CCGGTTCTGTCCTCGTACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGGCGAGCAG 482
Db      63871  CCGGTTCTGTCCTCGTACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGGCGAGCAG 63925
QY      483      GGAGACCCCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 542
Db      63926  GGAGACCCCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 63985
QY      543      GAGAGAGGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 602
Db      63986  GAGAGAGGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 64045
QY      603      GGCTCCAGGCGGAGGCGGGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
Db      64046  GGCTCCAGGCGGAGGCGGGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 64105
QY      663      GGGGATGAGTCTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 722
Db      64106  GGGGATGAGTCTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 64165
QY      723      CCGCGGATTCGATGCGATGCGCTCCGAGACGAGGCTCCCGGAGGAGGAGGAGGAG 782
Db      64166  CCGCGGATTCGATGCGATGCGCTCCGAGACGAGGCTCCCGGAGGAGGAGGAGGAG 64212
QY      783      GGAGCGCGCTGCGCTTCCAGTT 803
Db      64213  GGAGCGCGCTGCGCTTCCAGTT 64233
RESULT 7
AC126519/c      235390 bp      DNA      linear      HTG 20-NOV-2002
LOCUS      Rattus norvegicus clone CH230-159N5, WORKING DRAFT SEQUENCE.
DEFINITION
AC126519.3 GI:25138273
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 235390)
REFERENCE
Munry,D,Marle, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Albrooks,S, Amin,A, Angiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Bacs,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biewalo,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavaros,I, Cesar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,Y, Chen,Y, Chen,Z, Chu,J,
Claveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
David,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denison,S, Deramo,C, Ding,Y, Dink,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escoto,M, Eugene,C, Evans,C,A, Falle,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Georgievski,E, Geier,K, Gill,R, Grady,M, Guerra,M, Guetara,M,
Gunnarsson,P, Haaland,W, Hamill,C, Hamilton,N, Hernandez,J,
Harvey,Y, Havlik,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
Hollins,B, Howells,S, Huliyil,S, Hume,J, Idelbird,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kover,C, Liu,J,
Kovis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,M, Liu,Y, London,P, Longacre,S, Lopez,D,
Lorenzshewe,L, Louisedge,H, Lozano,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A,
Mangum,B, Mapua,P, Martin,K, Martin,K, Martinez,E,
Mawhiney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L,

```



```

OY      660  GGCGGCATGACGTCTGAGCTGGAGCGAGAGCCATGTCTGCCGACAGATGGCTCA 719
DB      26812 GGGCGGCATGACGTCTGAGCTGGAGCGAGAGCCATGTCTGCCGACAGATGGCTCA 26753
OY      720  GGACCCCGGATTCGATGCTCCCTCGAGACACGAGCTCCCGCAAGACGAGAGCAGA 779
DB      26752 GGACCCCGGATTCGATGCTCCCTCGAGACACGAGCTCCCGCAAGACGAGAGCAGA 26693
OY      780  CAAGAGCGCTCGCTTCCAGCTT 803
DB      26592 CAAGAGCGCTCGCTTCCAGCTT 26669

RESULT 8
AC125993/c
LOCUS   237695 bp  DNA  linear  HTG 10-MAY-2003
DEFINITION Rattus norvegicus clone CH230-74111, *** SEQUENCING IN PROGRESS
ACCESSION AC125993
VERSION   AC125993.3 GI:30522839
KEYWORDS  HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
          1 (bases 1 to 237695)
REFERENCE 1
AUTHORS   Murthy, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
          Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
          Arpaalbeck, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
          Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Bernheim, F.,
          Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
          Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
          Cadenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
          Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
          Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crease, A., D'Souza, L.,
          Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Decker, D.,
          Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
          Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
          Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
          Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
          Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, W.,
          Geiseczig, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M.,
          Gunatirne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K.,
          Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
          Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
          Hollins, B., Howell, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
          Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A.,
          Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
          Kowis, C., Kraft, C.L., Lepow, H., Lewin, J., Lewis, L., Li, Z., Liu, J.,
          Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
          Lorenz, L., Louisa, L., Louisa, H., Lozada, R.J., Lu, X., Ma, J.,
          Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A.,
          Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
          McWhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
          Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
          Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nair, L.,
          Nankevicius, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
          Nwaokwelen, O., Okunolu, G., Olarinmoyegun, A., Pal, S., Parks, K.,
          Pasternak, S., Paul, H., Perez, A., Perez, L., Plankoch, C.,
          Plopper, F., Polidexter, A., Popovic, D., Pitman, E., Pu, L.-L.,
          Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reith, R.,
          Rellly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
          Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
          Sanders, W., Savary, S., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
          Shetty, J., Shvartsberg, A., Sisson, I., Sitter, C.D., Smajic, D.,
          Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, D.,
          Steimle, N., Strong, R., Sutton, A., Svatek, A., Tabori, P., Taylor, C.,
          Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
          Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
          Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
          Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
          Wright, D., Wright, R., Wu, J., Yakub, S., Yen, T., Yoon, L., Yoon, V.,

```

```

TITLE     Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
          Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
          Weinstock, G. and Gibbs, R.A.
JOURNAL   Direct Submission
REFERENCE 2 (bases 1 to 237695)
AUTHORS   Unpublished
          Morley, K.C.
JOURNAL   Direct Submission
REFERENCE 3 (bases 1 to 237695)
AUTHORS   Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
          Rat Genome Sequencing Consortium.
          Direct Submission
          Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
          On May 10, 2003 this sequence version replaced gi:23096451.
          The sequence in this assembly is a combination of BAC based reads
          and whole genome shotgun sequencing reads assembled using Atlas
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
          in the feature table below represents a scaffold in the Atlas
          assembly (a 'contig-scaffold'). Within each contig-scaffold,
          individual sequence contigs are ordered and oriented. The sequence
          by sized gaps filled with Ns to the estimated size. The sequence
          may extend beyond the ends of the clone and there may be sequence
          contigs within a contig-scaffold that consist entirely of whole
          genome shotgun sequence reads. Both end sequences and whole genome
          shotgun sequence only contigs will be indicated in the feature
          table.

COMMENT   ----- Genome Center
          Center: Baylor College of Medicine
          Center code: BCM
          Web site: http://www.hgsc.bcm.tmc.edu/
          Contact: hgsc_help@bcm.tmc.edu
          ----- Project Information
          Center project name: G61Z
          Center clone name: CH230-74111
          ----- Summary Statistics
          Assembly program: Atlas 3.0
          Consensus quality: 229641 bases at least Q40
          Consensus quality: 231444 bases at least Q30
          Consensus quality: 232679 bases at least Q20
          Estimated insert size: 242786; sum-of-contigs estimation
          Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Bactinated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
      consists of 1 contigs. Gaps between the contigs
      are represented as runs of N. The order of the pieces
      is believed to be correct as given, however the sizes
      of the gaps between them are based on estimates that have
      been provided by the submitter.
      This sequence will be replaced
      by the finished sequence as soon as it is available and
      the accession number will be preserved.
      1 237695: contig of 237695 bp in length.
      Location/Qualifiers
          1. 237695
             /organism="Rattus norvegicus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10116"
             /clone="CH230-74111"
             1. 1995
             /note="wgs_end_extension
             clone_end:T7"
             complement(5394..6172)
             /note="clone_boundary
             clone_end:T7
             site:Ecort
             end_sequence:BH339391"
             complement(234275..235086)

```

misc_feature

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center Project name: KAMV
Center clone name: CH230-69F8

Summary Statistics

Assembly program: Phrap, version 0.990329
Consensus quality: 200303 bases at least Q40
Consensus quality: 206134 bases at least Q30
Consensus quality: 209370 bases at least Q20
Estimated insert size: 207464; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 4562: contig of 4562 bp in length
* 4563 4662: gap of unknown length
* 4663 259056: contig of 254394 bp in length
* 259057 259156: gap of unknown length
* 259157 260289: contig of 1143 bp in length
* 260300 260399: gap of unknown length
* 260400 262139: contig of 1740 bp in length.

FEATURES

source

1. 262139
/location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-69F8"
misc_feature
36039..38959
/note="wgs contig"
misc_feature
108494..109772
/note="wgs contig"
misc_feature
172942..174088
/note="wgs contig"
misc_feature
202428..204028
/note="wgs contig"

ORIGIN

Query Match 49.2%; Score 628; DB 2; Length 262139;
Best Local Similarity 88.1%; Pred. No. 4.6e-120;
Matches 709; Conservative 0; Mismatches 90; Indels 6; Gaps 2;

QY 3 GGGGGGAGGCGCGGAGCGACCGACCGTTCCTCCGCGAGCAGCTTCACCCCTGCGCCGA 62
Db 237331 GGGCGGAGGCGCGGAGCGACCGACCGTTCCTCCGCGAGCAGCGCCACCGACCGCGCA 237272
QY 63 TCCTTATCCGC---AGGCCACCAAGCGCGGAGTGGCTGAGGTTGGAGCCAGGGGCTG 119
Db 237271 TCCTTATCCGCACCGACCGCAAGCGCGGAGTGGCTGAGGTTGGAGCCAGGGGCTG 237212
QY 120 CCGACCGCGCGCCCTCTCTCTCCCGGCTACAGACAGCTATGCGCGCGAGTAGCT 179

Db 237211 CAGGCGGAGCGCCCTCTCTCTCCCGGCTACAGACAGTACTAGCCCGGAGTACTT 237152
QY 180 CGACAGCCACCAAGGGGCGACAGCTCATAGCCCTCTCTCTGAGATGGTCCCCGCTGCT 239
Db 237151 TGACAGCTATAGGAGGAGCGAGCTCATAGCCCTCTCTCTGAGATGGTCCCCGCTGCT 237092
QY 240 CAGAGCCGATACCTGCGGTGAGAGTAAACCCCGCGCGAGCGCTCGGTGAGTGTTC 299
Db 237091 CAGAGCGCGGACCGCTGCGGTGAGAGTAAACCCCGCGCGAGTGTTC 237032
QY 300 ACTGGGCGCGCGCGAGCGCTGAGAGTGTGAGGCTCGAGCCAGCCCGCGATGGG 359
Db 237031 GCTCGGCGCGCGCGAGCGCTGAGAGTGTGAGGCTCGAGCCAGCCCGCGATGGG 236972
QY 360 TTCTCTCAACCCCGGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 419
Db 236971 TTCTCTCAACCCCGGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 236912
QY 420 AGCCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
Db 236911 CGCCCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 236852
QY 480 CAGGAGACACCGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 539
Db 236851 CAGGAGACCGCGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 236792
QY 540 GCCGAGAGAGGTGCGCGCGAG 599
Db 236791 GCCGAGAGAGGTGCGCGAG 236732
QY 600 TCAGGCTGAGAGGAGAGGAGCGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
Db 236731 CGAGGCTGAGAGGAG 236675
QY 660 GCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 719
Db 236674 GCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 236615
QY 720 GAGACCCGCGTATTCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 779
Db 236614 GAGACCCGCGTATTCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 236555
QY 780 CAGAGAGCGCGCTGCGTTCAGAT 803
Db 236554 CAGAGAGCGCGCTGCGTTCAGAT 236531

RESULT 10

AC108848

LOCUS

DEFINITION

AC108848

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AC108848 212848 bp DNA linear HTG 28-Apr-2002
Mus musculus clone RP23-300K5, WORKING DRAFT SEQUENCE, 36 ordered
pieces.

AC108848 2 G1:20336129
HTG; HTG_PHS2; HTG_DRAFT; HTG_FULFILLTOP.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 212848)
Mus musculus, clone RP23-300K5

Unpublished

2 (bases 1 to 212848)

Batren, B., Lincon, L., Nuebaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Bouthagalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Coe, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mhova, T., Mwangi, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, J., Rietz, M., Riley, R., Riese, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
(bases 1 to 212848)

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, J., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Gange, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mhova, T., Mwangi, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, J., Rietz, M., Riley, R., Riese, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 28, 2002 this sequence version replaced gi:18450109.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Center code: W1BR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information
Center project name: L21014
Center clone name: 300_K5

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 200069 bases at least Q40
Consensus quality: 206302 bases at least Q30
Consensus quality: 208218 bases at least Q20
Insert size: 209348; sum-of-contigs
Quality coverage: 5.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

FEATURES

* the accession number will be preserved.
1 641: contig of 641 bp in length
642 741: gap of 100 bp
742 1427: contig of 686 bp in length
1428 1527: gap of 100 bp
1528 3170: contig of 1643 bp in length
3171 3270: gap of 100 bp
3271 4396: contig of 1126 bp in length
4397 4497: gap of 100 bp
4497 5909: contig of 1413 bp in length
5910 6010: gap of 100 bp
6010 7583: contig of 1574 bp in length
7584 7683: gap of 100 bp
7684 8975: contig of 1292 bp in length
8976 9075: gap of 100 bp
9076 10257: contig of 1182 bp in length
10258 10358: gap of 100 bp
10358 11608: contig of 1251 bp in length
11609 11708: gap of 100 bp
11709 13289: contig of 1581 bp in length
13290 13389: gap of 100 bp
13390 15335: contig of 1946 bp in length
15336 15435: gap of 100 bp
15436 16250: contig of 815 bp in length
16251 16350: gap of 100 bp
16351 17669: contig of 1319 bp in length
17670 17769: gap of 100 bp
17770 19907: contig of 2138 bp in length
19908 20007: gap of 100 bp
20008 21471: contig of 1463 bp in length
21471 21571: gap of 100 bp
21571 23086: contig of 1516 bp in length
23087 23186: gap of 100 bp
23187 24723: contig of 1537 bp in length
24724 24823: gap of 100 bp
24823 27746: contig of 2923 bp in length
27747 27846: gap of 100 bp
27847 29873: contig of 2027 bp in length
29873 29973: gap of 100 bp
29974 33212: contig of 3239 bp in length
33213 33312: gap of 100 bp
33313 61733: contig of 28420 bp in length
61733 61833: gap of 100 bp
61833 63793: contig of 1960 bp in length
63793 63893: gap of 100 bp
63893 67232: contig of 3339 bp in length
67232 67331: gap of 100 bp
67331 70922: contig of 3591 bp in length
70922 71023: gap of 100 bp
71023 76930: contig of 5908 bp in length
76931 77030: gap of 100 bp
77031 82623: contig of 553 bp in length
82623 82723: gap of 100 bp
82723 87875: contig of 5152 bp in length
87876 87975: gap of 100 bp
87976 93774: contig of 5799 bp in length
93775 93874: gap of 100 bp
93875 104205: contig of 10331 bp in length
104206 104305: gap of 100 bp
104306 119104: contig of 14789 bp in length
119105 119204: gap of 100 bp
119205 132299: contig of 13095 bp in length
132300 132399: gap of 100 bp
132400 148970: contig of 16571 bp in length
148971 149070: gap of 100 bp
149071 155714: contig of 16644 bp in length
155715 156814: gap of 100 bp
156815 163368: contig of 17355 bp in length
163369 183470: gap of 100 bp
183471 210291: contig of 26822 bp in length
210292 210391: gap of 100 bp
210392 212848: contig of 2457 bp in length.

Location/Qualifiers

*	* is believed to be correct as given, however the sizes
*	* of the gaps between them are based on estimates that have
*	* provided by the submittor.
*	* This sequence will be replaced
*	* by the finished sequence as soon as it is available and
*	* the accession number will be preserved.
*	* 641: contig of 641 bp in length
*	* 1
*	* 642 741: gap of 100 bp
*	* 742 1477: contig of 686 bp in length
*	* 1428 1577: gap of 100 bp
*	* 1528 3170: contig of 1643 bp in length
*	* 3171 3270: gap of 100 bp
*	* 3271 4336: contig of 1126 bp in length
*	* 4397 4436: gap of 100 bp
*	* 4497 5909: contig of 1413 bp in length
*	* 5910 6009: gap of 100 bp
*	* 6010 7533: contig of 1574 bp in length
*	* 7564 7584: gap of 100 bp
*	* 7684 8975: contig of 1292 bp in length
*	* 8776 9075: gap of 100 bp
*	* 9076 10257: contig of 1182 bp in length
*	* 10258 103577: gap of 100 bp
*	* 10358 11608: contig of 1251 bp in length
*	* 11609 11708: gap of 100 bp
*	* 11709 13289: contig of 1581 bp in length
*	* 13290 13389: gap of 100 bp
*	* 13390 15335: contig of 1946 bp in length
*	* 15336 15435: gap of 100 bp
*	* 15436 16230: contig of 815 bp in length
*	* 16251 16350: gap of 100 bp
*	* 16351 17669: contig of 1319 bp in length
*	* 17670 17769: gap of 100 bp
*	* 17770 19977: contig of 2136 bp in length
*	* 19908 200077: gap of 100 bp
*	* 20008 21470: contig of 1463 bp in length
*	* 21471 21570: gap of 100 bp
*	* 21571 23086: contig of 1516 bp in length
*	* 23087 23186: gap of 100 bp
*	* 23187 24733: contig of 1537 bp in length
*	* 24724 24823: gap of 100 bp
*	* 24824 27766: contig of 2923 bp in length
*	* 27747 27866: gap of 100 bp
*	* 27847 298973: contig of 2027 bp in length
*	* 29874 29973: gap of 100 bp
*	* 29974 33242: contig of 3235 bp in length
*	* 33213 33312: gap of 100 bp
*	* 33313 61732: contig of 28440 bp in length
*	* 61733 61832: gap of 100 bp
*	* 61833 63792: contig of 1960 bp in length
*	* 63793 63892: gap of 100 bp
*	* 63893 67321: contig of 3339 bp in length
*	* 67332 70922: gap of 100 bp
*	* 70923 71022: gap of 100 bp
*	* 71023 76930: contig of 5908 bp in length
*	* 76931 77030: gap of 100 bp
*	* 77031 82633: contig of 5593 bp in length
*	* 82634 82723: gap of 100 bp
*	* 82724 87975: contig of 5152 bp in length
*	* 87976 93776: gap of 100 bp
*	* 93777 93874: contig of 5799 bp in length
*	* 93875 104205: gap of 100 bp
*	* 104206 104305: contig of 10331 bp in length
*	* 104306 119104: contig of 14799 bp in length
*	* 119105 119204: gap of 100 bp
*	* 119205 132289: contig of 13055 bp in length
*	* 132300 132389: gap of 100 bp
*	* 132900 148970: contig of 16571 bp in length
*	* 148971 149070: gap of 100 bp
*	* 149071 165714: contig of 16644 bp in length
*	* 165715 165814: gap of 100 bp
*	* 165815 183369: contig of 17555 bp in length

[illegible]

RESULT	13			
LOCUS	AY283176			
DEFINITION	Xenopus laevis zygote arrest 1 (Zarl) mRNA, complete cds.	1052 bp	mRNA	linear
ACCESSION	AY283176			
VERSION	AY283176.1	GI:30908934		
KEYWORDS				
SOURCE				
ORGANISM	Xenopus laevis (African clawed frog)			
	Xenopus laevis			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Amphibia; Batrachia; Anura; Mesobatrachina; Pipoidae; Pipidae;			
	Xenopodinae; Xenopus.			
REFERENCE	1 (bases 1 to 1052)			
AUTHORS	Wu,X., Wang,P., Brown,C.A., Zilinski,C.A. and Matzuk,M.M.			
TITLE	Zygote arrest 1 (Zarl) is an evolutionarily conserved gene expressed in vertebrate ovaries			
JOURNAL	Biol. Reprod. 69 (3), 861-867 (2003)			
MEDLINE	22811438			
PubMed	12773403			
REFERENCE	2 (bases 1 to 1052)			
AUTHORS	Wu,X. and Matzuk,M.M.			
TITLE	Direct Submission			
JOURNAL	Submitted (24-Apr-2003) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA			
FEATURES				
Source	1..1052			
	Location/Qualifiers			

```

gene
CDS
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
1..1052
/gene="zari"
24..911
/gene="zari"
/note="maternal factor"
/codon_start=1
/product="zygote arrest 1"
/protein_id="AAP37038.1"
/db_xref="GI:30908935"
/translation="MIPAINSYRYLLENFNKNGSMKQNTLASIGTGDYCDNYQAPR
QKRLISGVNENLPLICRANTRDVGAVNRRQASVQCSIGPTLLRRRGAIRKPP
PEQSPASPTKTVRPFRTIAYSVFVAAGRKLAPFDEGVNLEKEEAVSESEEGRDE
GKQDGEIKKEGMKMDKTDDEEAPADQRRPKQFLEQKRGYGHCDNDIRMSASVYWC
GENTNKKYFKQCRICQKSYNPEYVEDIMCQSCKQTRCACPYKLAHVDKPRHODLOG
RCKGRKRLCSDSTSFKKYII"

```

	Query Match	16.1%	Score 205.8 ; DB 5 ; Length 1052 ;
	Best Local Similarity	80.1% ; Pred. No. 2.4e-32 ;	
	Matches	254 ; Conservative 0 ; Mismatches 62 ; Indels 1 ; Gaps 1 ;	
QY	796	TTCCAGTTCCTTAGACGACAAAGTAACGGCTACTATCACTGCACAAGA CTGCCAATCCCGGTGG	855
Dd	594	TTCAGTTCCTCGAGACGAAGTAACGATATTATCACTGTAGAAGCTGCACATCCCGCTGG	653
QY	856	GAGACCGCTATGATGTGTGTGTGCAGGGCACAGATAAGTGTTACTTCAAACAGTTCTG	915
Dd	654	GAGACCGCTATGATGTGTGTGTGCAGGAACCAATAAGGTG-TACTTCAAACAGTTCTG	712
QY	916	CCGAGTGTGTAGAAATCCTACAACCCTTACAGAGTGGAGACATCACTGTCAAAGTTG	975
Dd	713	CAGACATGTGAGAAATCTATTAATCCCTACCGTGTGGAAACATCATGTGTCAAGCTG	772
QY	976	TAAAAGAACTTGAATGTGCTCTGCCAAGTCAGATTTCGCCAGTGGACCTTAACGCCCCCA	1035
Dd	773	CAACCAAGCGGAATGCGCGTGTCTGTCCAATCAGTACAGTTAACCCCAAGAGGCCCA	832
QY	1036	TCGCGAAGACTTGTGTGGAGATCAAGACAACAAAGCCTGTCTCGACAGACACTTTCAG	1095
Dd	833	CCGCGAGATGTGTGTGGAGATCAAGACAACAAAGCCTGTCTGTGTACAGACACTTTTAC	892
QY	1096	CTTCAATATCATCATTT 1112	

```

DB              893  CTTCAGATATCATTT  909

RESULT 14
LOCUS          AY283177
DEFINITION     Takifugu rubripes zygote arrest 1 (Zar1) mRNA, complete cds.
ACCESSION      AY283177
VERSION        AY283177.1  GI:30908936
KEYWORDS
SOURCE
ORGANISM       Takifugu rubripes (Fugu rubripes)
                Takifugu rubripes
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Euteleostei;
                Acanthopterygii; Acanthopterygii; Perciformes;
                Tetraodontiformes;
                Tetraodontidae; Tetraodontidae; Takifugu.
                1 (bases 1 to 963)
                Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matczuk, M.M.
                Zygote arrest 1 (Zar1) is an evolutionarily conserved gene
                expressed in vertebrate ovaries
                Biol. Reprod. 69 (3), 861-867 (2003)
                2811438
                JOURNAL
                PUBMED 1273403
                REFERENCE
                2 (bases 1 to 963)
                Wu, X. and Matczuk, M.M.
                Direct Submission
                Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
                location/Qualifiers
                1..963
FEATURES
SOURCE

```

```

/organism="Iaktiungu ludripes"
/mol_type="mRNA"
/db_xref="taxon:31033"
1..963
/gene="Zar1"
1..963
/gene="Zar1"
/note="maternal factor"
/codon_start=1
/product="zygote arrest 1"
/protein_id="Aap37039.1"
/db_xref="GI:30908937"
/translation="MATYCDPEVDSDYFSSYNPNGRYPHRDAGMKYKSLVSHGDDT
SAFSSNOORAKLSILSQINPKLIPRLRKANTDVAAYOVAPKSDASVGSIGPTLLIV
VKEELIRRLNTPGPGTPOKTEGGEVRYPTLLVSLIARSVTSFLVNGCKDPAHE
AQCEELIPGEPPQKCGENCGAGETNNALPQGRQPSQSDAOTPADABESKKKARVFPQ
LEKRIYTHCRNCRNWSAIVWCVOGDTNRYVTRKQRCQCKQNPVREVDITCVON
KACACACETPRHVDPKRPHRQDLCGKCGKGRSLCSFTSKYIV"

```

	Query Match	Best Local Similarity	15.3%	Score 19.4	DB 5	Length 963
	Matches	255	Conservative	0	Mismatches	82
					Indels	1
					Gaps	1
Qy	776	AGGACAAAGAGGCGCCCTGCTTCCTTCCAGTTCTTTAAGACGAAAGTACGGCTACTATATACCTGCA	835			
Dh	626	AGGCGAAAGGCGCGTGTCCGCTTCCAGTTCTTTGGAACGAAAGTACGGCTACTATATACCTGCA	685			
Qy	836	AGGACTGCAAAATCCGGGTGGAGAGCGCTATGTGTGTGTCAGAGGACCAAGTAAAG	895			
Dh	686	GAGAAATCAACCTGCGATGGGAAGAGCGCGTACGTTTGTGTGCTTCAAGGCACTTAAACAGG	745			
Qy	896	TGTTTACTTCAAACAGTTTTCGCCAGTGTGTGAAGAAATCTTACAAACCTTACAGAGTGGAG	955			
Dh	746	T-TTAACTCAAGACGTTCTGTGAAGAAATGCAAAAAGACTTAAACCGTACCGGTGAG	804			
Qy	956	GACATCAACCTGTCAAAGTTGTAAAGAACTAGATGTGCTGCCAGTCAAGATTTGCCAC	1015			
Dh	805	GACATCAACGTACAGTATGCAACAAGGCCCGGTGTGTCCTGCGCAAGAAAGCAGCGCCAC	864			
Qy	1016	GTCGACCTTAAAGCCCCCTTCGGCAAGACTTGTGTGGAGATGCAAGACAAACGCGTG	1075			
Dh	865	GTTGACCCCAAGAGGCCCCACAGACGACGAACTGTGTGCGCAGGTGCAAGGGCCAAAGCGGCTG	924			

Db	121	CGACCCGCGCCCCCTCTTCCTCCCGGCGTACAGACACTCATGGCCGCGAGTACGTC	18
OY	181	GACAGCGACACCAAGGGGACAGCTCATAGGCCCTGCTGCGGGATGGATCCCGGTCCGTC	240
Db	181	GACAGCGACACAGGGGACAGCTCATAGGCCCTGCTGCGGGATGGATCCCGGTCCGTC	240
OY	241	AGCAGCCGTGACGCTGCGGTGACAGTGAACCCGCGCGAGCGCTCGGTGACAGTTC	300
Db	241	AGCAGCCGTGACGCTGCGGTGACAGTGAACCCGCGCGAGCGCTCGGTGACAGTTC	300
OY	301	CTCGGCGCGCGACGCTGAGCGCTGAGGGGTGCAGGCCAGGCCCGGATCCGGT	360
Db	301	CTCGGCGCGCGACGCTGAGCGCTGAGGGGTGCAGGCCAGGCCCGGATCCGGT	360
OY	361	TCTGTCAAACCCCGTGGACGCGCGCGCGGAGATCCCGCGATCTCGGACAGCCGTA	420
Db	361	TCTGTCAAACCCCGTGGACGCGCGCGCGGAGATCCCGCGATCTCGGACAGCCGTA	420
OY	421	GCCCCGTTCTGTCCTGCTGACCTTTCTGTGGCTCTCTCTCACTGAGGTTCGGAGGC	480
Db	421	GCCCCGTTCTGTCCTGCTGACCTTTCTGTGGCTCTCTCTCACTGAGGTTCGGAGGC	480
OY	481	AGGCAGACACCCACGAGGAGAGGGAGACCCCGCATCTCTGGAGACCCGAGAACCGAG	540
Db	481	AGGCAGACACCCACGAGGAGAGGGAGACCCCGCATCTCTGGAGACCCGAGAACCGAG	540
OY	541	CCGAGAGAGGTGGCCGCGAGGAAAGCGTCCCGACGCGGAGAGGAGAGGGCGCATGTT	600
Db	541	CCGAGAGAGGTGGCCGCGAGGAAAGCGTCCCGACGCGGAGAGGAGAGGGCGCATGTT	600
OY	601	CAGGCTGCAGGGCAGGCGCGGTGGAGACAGCCACACCCGAGAGCCGGAACAGTGTG	660
Db	601	CAGGCTGCAGGGCAGGCGCGGTGGAGACAGCCACACCCGAGAGCCGGAACAGTGTG	660
OY	661	GCGGCGATGCACTCTGAGCCTTGAGGCGGAGGAGCCATGCTCGCGGAGAGTGGCTAG	720
Db	661	GCGGCGATGCACTCTGAGCCTTGAGGCGGAGGAGCCATGCTCGCGGAGAGTGGCTAG	720
OY	721	GAACCCGGTGATTGGAGTACCCCTCTGAGACAGAGCCTTCCCGGAAAGACGAGACGAGAC	780
Db	721	GAACCCGGTGATTGGAGTACCCCTCTGAGACAGAGCCTTCCCGGAAAGACGAGACGAGAC	780
OY	781	AAGAGCGCGCTGCGTTTCAGTTCCTTAGAGCGAGAGTACGGCTATCATCTGCAAGGAC	840
Db	781	AAGAGCGCGCTGCGTTTCAGTTCCTTAGAGCGAGAGTACGGCTATCATCTGCAAGGAC	840
OY	841	TGCAAAATCCGCTGGGAGAGCGCTATGTGTGTGTGCAAGGACCAAGTAAAGGTGTTA	900
Db	841	TGCAAAATCCGCTGGGAGAGCGCTATGTGTGTGTGCAAGGACCAAGTAAAGGTGTTA	900
OY	901	CTTCAAAACAGTTCCTGCGAGTGTGTGAGAAATCTCAACAACCTTCACAGGTGGAAGAT	960
Db	901	CTTCAAAACAGTTCCTGCGAGTGTGTGAGAAATCTCAACAACCTTCACAGGTGGAAGAT	960
OY	961	CACCTGTCAAAAGTTGTAAAGAACTAGATGTGCTGTCCCGATGATTTCCGACGTGGA	1020
Db	961	CACCTGTCAAAAGTTGTAAAGAACTAGATGTGCTGTCCCGATGATTTCCGACGTGGA	1020
OY	1021	CCCTTAAACGCCCCCATTCGGCAAGATTTGTGTGGAGATGCAAGGACAAACGCTTCCTG	1080
Db	1021	CCCTTAAACGCCCCCATTCGGCAAGATTTGTGTGGAGATGCAAGGACAAACGCTTCCTG	1080
OY	1081	CGACAGCACCTTCAGCTTCAAATACATCAATTTAGTGAAGTGAAGAAAGTTCGTGAGA	1140
Db	1081	CGACAGCACCTTCAGCTTCAAATACATCAATTTAGTGAAGTGAAGAAAGTTCGTGAGA	1140
OY	1141	TGGGCGTATGGAATGAGCAAGTGAAGTTTCGCCCTTCACCTCTCCCTTCCCAAT	1200
Db	1141	TGGGCGTATGGAATGAGCAAGTGAAGTTTCGCCCTTCACCTCTCCCTTCCCAAT	1200
OY	1201	TCTTCATGACAGACAGTGTACTTGTGATATAAGCTGTGAATTAAGGTATTCGAACA	1260
Db	1201	TCTTCATGACAGACAGTGTACTTGTGATATAAGCTGTGAATTAAGGTATTCGAACA	1260

[illegible]

QY 1205 CATGACGACGAGTGTACTTGATATATAAGCTGTGATATAAGTATTGCAACA 1260
DB 250 CATGACGACGAGTGTACTTGATATATAAGCTGTGATATAAGTATTGCAACA 305

RESULT 5

US-09-844-864-25
Sequence 25, Application US/09844864
Patent No. US20020042926A1
GENERAL INFORMATION:
APPLICANT: Matzuk, Martin
APPLICANT: Ren, Yongsheung
APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106,020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 25
LENGTH: 375
TYPE: DNA
ORGANISM: mus musculus
US-09-844-864-25

Query Match

Best Local Similarity 20.4%; Score 260.6; DB 9; Length 375;
Matches 266; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 986 AGATGCTGCTCCCGAGTCAGATTTCGCCACGTGAGACCTTAAAGCCCCATGCGCAAGC 1045
DB 101 AGATGCTGCTCCCGAGTCAGATTTCGCCACGTGAGACCTTAAAGCCCCATGCGCAAGC 160
QY 1046 TTGTGTGGGAGATGCAAGGCAAGGCTGTCTGTGAGAGAGACCTTCAAGTCAAAATAC 1105
DB 161 TTGTGTGGGAGATGCAAGGCAAGGCTGTCTGTGAGAGAGACCTTCAAGTCAAAATAC 220
QY 1106 ATCATTTAGTAGAGAGTGAAGAAAGCTTCTGCTAGATGGGCTAATGGAATGACAAGTGA 1165
DB 221 ATCATTTAGTAGAGAGTGAAGAAAGCTTCTGCTAGATGGGCTAATGGAATGACAAGTGA 280
QY 1166 GCTTCTCCCTTTCACCTCTTCCCTTCCAAATCTTCATGACGAGAGAGTACTTG 1225
DB 281 GCTTCTCCCTTTCACCTCTTCCCTTCCAAATCTTCATGACGAGAGAGTACTTG 340
QY 1226 GATATAAGCCTGTGATATAAGTATTGCAACA 1260
DB 341 GATATAAGCCTGTGATATAAGTATTGCAACA 375

RESULT 6

US-09-844-864-19
Sequence 19, Application US/09844864
Patent No. US20020042926A1
GENERAL INFORMATION:
APPLICANT: Matzuk, Martin
APPLICANT: Ren, Yongsheung
APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106,020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0

/ SEQ ID NO 19
/ LENGTH: 123
/ TYPE: DNA
/ ORGANISM: mus musculus
US-09-844-864-19

Query Match 7.7%; Score 98.8; DB 9; Length 123;
Best Local Similarity 98.0%; Pred. No. 7.8e-19;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 795 TTTCAGTTCTTAGAGAGAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 854
DB 9 TTTCAGTTCTTAGAGAGAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 68
QY 855 GGAAGGCGCCTATGTGTGTGTGTGACAGGACCCAGTAAGT 896
DB 69 GGAAGGCGCCTATGTGTGTGTGTGACAGGACCCAGTAAGT 110

RESULT 7

US-09-844-864-23
Sequence 23, Application US/09844864
Patent No. US20020042926A1
GENERAL INFORMATION:
APPLICANT: Matzuk, Martin
APPLICANT: Ren, Yongsheung
APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106,020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 23
LENGTH: 123
TYPE: DNA
ORGANISM: mus musculus
US-09-844-864-23

Query Match 7.7%; Score 98.8; DB 9; Length 123;
Best Local Similarity 98.0%; Pred. No. 7.8e-19;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 795 TTTCAGTTCTTAGAGAGAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 854
DB 9 TTTCAGTTCTTAGAGAGAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 68
QY 855 GGAAGGCGCCTATGTGTGTGTGTGACAGGACCCAGTAAGT 896
DB 69 GGAAGGCGCCTATGTGTGTGTGTGACAGGACCCAGTAAGT 110

RESULT 8

US-09-844-864-20
Sequence 20, Application US/09844864
Patent No. US20020042926A1
GENERAL INFORMATION:
APPLICANT: Matzuk, Martin
APPLICANT: Ren, Yongsheung
APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106,020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 105
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-20

Query Match
Best Local Similarity 100.0%; Pred. No. 5.7e-11; Length 105;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 899 TACTTCAACAGTTCCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 958
Db 19 TACTTCAACAGTTCCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 78

QY 959 ATCAGCTGTCAA 970
Db 79 ATCAGCTGTCAA 90

RESULT 9
US-09-844-864-24
; Sequence 24, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheung
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 105
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-24

Query Match
Best Local Similarity 98.6%; Pred. No. 1.7e-10; Length 105;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 899 TACTTCAACAGTTCCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 958
Db 19 TACTTCAACAGTTCCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 78

QY 959 ATCAGCTGTCAA 970
Db 79 GTCAGCTGTCAA 90

RESULT 10
US-10-412-699B-1133
; Sequence 1133, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Ficomm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Biron, Pierre E.
; APPLICANT: Pineda, Omalra
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
```

```

; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubeil, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumamoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: M01-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1133
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-412-699B-1133

Query Match
Best Local Similarity 43.4%; Pred. No. 3.7e-05; Length 1500;
Matches 251; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 137 CCTTCTCTCCCGGCTACAGACATGAGCCCGCGAGTACGTGACAGCCACAGCGGG 196
Db 298 CCTTCTTACACACCCCGACAGAGCTCTCGAAGAGGAGTACTACAGACAGAGCTCCCGG 357

QY 197 CACAGCTATGAGCCCTGTGTGCGGAGTGGGTCCCGGTCGTGAGAGCCGTGAGCGTG 256
Db 358 AGAAGAGCGGCGCTTACAGCGCGGAGCAGAGTGTCTGTGAGAGAGCTTGAAGAG 417

QY 257 CGGTGAGGTGAACCCCGCGCGGACCGCTCGTGTGAGTGTCACTCGGCGCGACGC 316
Db 418 AGAACAAGCTGAGACCGGAGCGGAAGAAGAGCTGGCGGAGAGCTGAGCGCGC 477

QY 317 TGCAGCTGTGAGGAGTGGCGAGGCGCCCGACAGCGCGGATCGGGTCTGTCAACCCCGTG 376
Db 478 GCGAGGTGCGGTGTGTGTTCCAGAACCGCGCGCGCGCTGGAAGACCAAGACGCTTCAAGC 537

QY 377 GCCAGCGCGGCGCGGAGATCCCGGAGTCTTGGAGACCGTGAAGCCCGCTTCTGCTCG 436
Db 538 GCGACTTGAACCGCTCAAGGCGTGTTCGACGCGCTCCGCGCGGACACAGACCGCTTCC 597

QY 437 TGACCTTGTGTGCTCTCTCTCTCTCACTGAGAGTTGGGAGAGGAGGACAGACACCCAG 496
Db 598 TCAGAGCAACCCAGCGCTTCACTTCAAGTGTGATGTGTGAACGAGAAAGTGTCAAGAGA 657

QY 497 AGGAGAGGAGGAGCGCGCATCTCGGAGAACCCCGGAGACCGGAGCCGAGAGAGGTGCGG 556
Db 658 AGGAGAGGAGGAGCGCGCATCTCGGAGAACCCCGGAGCGCGGTGAGAGTCCCGGCTTGCCTG 717
```

QY 557 CGAGAAAGCGTCCCCAGCCGGAAGAGAGCGATGTTCAAGCTGACAGGCGAG 616
DB 718 CGCGCGAGTGAAGTGGCGGTCCCGGACCGGAGAGACCGGCGCTGAGAGAGCGCGCG 777
QY 617 CCGGTGGAGAC 676
DB 778 CGGCTTCGAGAGAC 837
QY 677 AGCCTGGAGAGCGAGACACACACACACACACACACACACACACACACACAC 715
DB 838 GCGCGGAGAGCGCGTGTGTGACACGAGACCGGACACTGG 876

RESULT 11
US-10-374-780A-680
; Sequence 680, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Racciliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Biron, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omara
; APPLICANT: Yu, Guo-Liang
; TITLE OR INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 680
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: *Oryza sativa*
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G409
US-10-374-780A-680

Query Match 4.2%; Score 54.2; DB 15; Length 1500;
Best Local Similarity 4.4%; Pred. No. 3,7e-05;
Matches 251; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 137 CCTCCCTCCCGGCTACAGACAGTCTGCGCGGAGTCTGTCAGACGACACACAGG 196
DB 298 CTTCTTTCACACACCGCGAGAGCTCTGAGAGAGAGTCTGACGAGACAGTCTCCG 357

QY 197 CACAGCTCATGAGCCCTGCTGCGGATGGATGCCGCTCCCGTCCGTCAGACGCCGTG 256
DB 358 AGAAGAACGGCGGCTCAGCGCGGAGAGAGTGTGATCTGTGAGAGAGAGCTTCAGAGAG 417
QY 257 CCGTTCAGGTAAACCGCGCGCGAGACGCTCTGTGTGAGTGTTCACCTGAGGCGCGAC 316
DB 418 AGAACAGCTGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
QY 317 TGCAGCTTCAGAGGTGCGGAGACGAGCCCGAGAGCGGATTCGATTCAGACCCGCTG 376
DB 478 GCGAGGTGCGGTGTGTGTTCAGAACCGCCCGCGCGCTGAGAACAGAGAGTTCAGAGC 537
QY 377 GCGACCGCGCGCGGAGATCCCGGATCTCTGAGACCGTTCAGACCCGCTTCGTCG 436
DB 538 GCGACTTCAGACCGGCTCAGAGCGGTGTTCAGACCGCTTCGCGAGACACACACCGCTCC 597
QY 437 TGAAGCTTCAGAGCGCTCTCTCTCACTGAGAGTTCGAGAGAGAGAGAGAGAGAGAG 496
DB 598 TCCAGAGAACACCGCTCTCACTCTCACTGATGTCTTGAACGAGAGAGTTCAGAGAG 657
QY 497 AGGAGAGAGAGAGCGCGGATCTCTGAGAGACCGGAGACCGAGACCGAGAGAGTTCG 556
DB 658 AGGAGACAGACAG 717
QY 557 CGAGAAAGCGGTCCCCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
DB 718 CGGCGGAGCTAAGGTCCCTCCCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
QY 617 CCGGTGGAGACAGAGACACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 676
DB 778 CGGCTTCGAG 837
QY 677 AGCCTGGAG 715
DB 838 GCGCGGAGAGCGCGTGTGTGACAGAGACCGGACACTGG 876

RESULT 12
US-10-084-846A-99
; Sequence 99, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUEHLENER, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTOLD, ANDREAS
; TITLE OR INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 10. 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 99
; LENGTH: 987
; TYPE: DNA
; ORGANISM: *Streptomyces viridochromogenes*
; FEATURE:
; OTHER INFORMATION: aviga data: partial sequence of coding strand 1; nucleotide 1
; OTHER INFORMATION: corresponds to nucleotide 45,341 of coding strand 1.
US-10-084-846A-99

Query Match 4.0%; Score 51.6; DB 15; Length 987;
Best Local Similarity 4.6%; Pred. No. 0.00017;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 80 CCAAGCGCGGAGTGTGTGAGTTCGAGAGCAGAGGCTGCGGACCGCGCGCGCTCCT 139
DB 260 CCGCGGTGCGAGCGCGCTTCAGAGACCTTGACACACAGCGTTCGAGACCGCGCGCT 319

```
QY 140 TCCTCCCGGCTACAGACAGCTGAGCGCGAGTACGTGACAGCCACAGCGGAC 199
DB 320 TCAGACAGCTTTGGCGCGGACCTTCTCGGTACCTGGGAGAGACCCGACCTGCT 319
QY 200 AGCTCATGGCCCTCTGCTCGCGATGGGTCCCGGTGCTGACAGCCGCTGACGCTGG 259
DB 380 GGCTCTACAAAGCGCGGATGAGCCAGGACCGGTGGAAATCGCGGCTGTGCGCGCG 439
QY 260 TGACAGTGAACCCGCGCGACGCTGCTGCTGACGTGCTTCACTCGGCGCGACGCTGC 319
DB 440 ACCAGAGCTTCTCCGGCGCTCCGACCGTGTGAGACGTGCGGCGCGGAGACGCTGC 499
QY 320 AGCTGCAAGGTGCCGAGCCAGCCCC 345
DB 500 TGGCGCGGTGCTGCGCGGACCC 525

RESULT 13
US-10-084-846A-1
; Sequence 1, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLLENWEG, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084, 846A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 59816
; TYPE: DNA
; ORGANISM: Streptomyces viridochromogenes
US-10-084-846A-1

Query Match 4.0%; Score 51.6; DB 15; Length 59816;
Best Local Similarity 49.6%; Pred. No. 0.0013;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 80 CCAAGCCGGGAGTGGCTGAGAGTTGGAGCCAGGGCTGCCAGCCGCCCTCTCT 139
DB 45600 CCGCGGTGCGACGCGCTGGAGACCTTGACCAACGCTCCGACCGCGCGGCT 45659
QY 140 TCCTCCCGGCTACAGACAGCTGAGCGCGGAGTACGTGACAGCCACAGCGGAC 199
DB 45660 TCAGACAGCTTTGGCGCGGACCTTCTCGGTACCTGGGAGAGACCCGACCTGCT 45719
QY 200 AGCTCATGGCCCTCTGCTCGCGATGGGTCCCGGTGCTGACAGCCGCTGACGCTGG 259
DB 45720 GGCTCTACAAAGCGCGGATGAGCCAGGACCGGTGGAAATCGCGGCTGTGCGCG 45779
QY 260 TGACAGTGAACCCGCGCGACGCTGCTGCTGACGTGCTTCACTCGGCGCGACGCTGC 319
DB 45780 ACCAGAGCTTCTCCGGCGCTCCGACCGTGTGAGACGTGCGGCGCGGAGACGCTGC 45839
QY 320 AGCTGCAAGGTGCCGAGCCAGCCCC 345
DB 45840 TGGCGCGGTGCTGCGCGGACCC 45865

RESULT 14
US-10-084-846A-2/c
; Sequence 2, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
```

```
APPLICANT: MUHLLENWEG, AGNES
APPLICANT: TREPZER, AXEL
APPLICANT: BECHTHOLD, ANDREAS
TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
FILE REFERENCE: 1974-005
CURRENT APPLICATION NUMBER: US/10/084, 846A
PRIOR FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: PCT/EP01/09815
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: DE 101 09 166.4
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 2
LENGTH: 59816
TYPE: DNA
ORGANISM: Streptomyces viridochromogenes
US-10-084-846A-2

Query Match 4.0%; Score 51.6; DB 15; Length 59816;
Best Local Similarity 49.6%; Pred. No. 0.0013;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 80 CCAAGCCGGGAGTGGCTGAGAGTTGGAGCCAGGGCTGCCAGCCGCCCTCTCT 139
DB 14217 CCGCGGTGCGACGCGCTGGAGACCTTGACCAACGCTCCGACCGCGGCT 14158
QY 140 TCCTCCCGGCTACAGACAGCTGAGCGCGGAGTACGTGACAGCCACAGCGGAC 199
DB 14157 TCAGACAGCTTTGGCGCGGACCTTCTCGGTACCTGGGAGAGACCCGACCTGCT 14098
QY 200 AGCTCATGGCCCTCTGCTCGCGATGGGTCCCGGTGCTGACAGCCGCTGACGCTGG 259
DB 14097 GGCTCTACAAAGCGCGGATGAGCCAGGACCGGTGGAAATCGCGGCTGTGCGCG 14038
QY 260 TGACAGTGAACCCGCGCGACGCTGCTGCTGACGTGCTTCACTCGGCGCGACGCTGC 319
DB 14037 ACCAGAGCTTCTCCGGCGCTCCGACCGTGTGAGACGTGCGGCGCGGAGACGCTGC 13978
QY 320 AGCTGCAAGGTGCCGAGCCAGCCCC 345
DB 13977 TGGCGCGGTGCTGCGCGGACCC 13952

RESULT 15
US-10-156-761-6879
; Sequence 6879, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6879
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; NAME/KEY: CDS
; LOCATION: (1)..(1185)
US-10-156-761-6879
```

Query March 4.0%; Score 51; DB 14; Length 1185;
Best Local Similarity 45.0%; Pred. No. 0.00029;
Matches 230; Conservative 0; Mismatches 280; Indels 1; Gaps 1;

QY 182 ACAGCCACGAGCGGACAGCTATGACCTGCTGTGCGAGTGGGTCCCGGTGCA 241
DB 503 ACCGCGCGCTGTGTGCGCGCCCTCCAGGAGGAGCTCCCGGCGCGGTGACCGTCGAGAAAG 562
QY 242 GCAGCCCTGACGCTGCGGTGAGTGAACCCGCGCGGACGCGCTGGTGCAGTTCAC 301
DB 563 AGACCAACTGTGCGCGCTGTGCGGACAGCGAGGCGCGCGCGGACCGGACAGT 622
QY 302 TCGGCGCGCGACGCTGACGCTGACAGGTGCGGACGCGGACCGCGGATCGGATT 361
DB 623 TCGT 682
QY 362 CTTGTCAACCCGTTGCGGACCGCGCGCGGAGATCCCGCGATCTGTGCGAGCCGTAG 421
DB 683 GCGGTGCGCGCTGTGCGCGGACCGCGGAGATCGGCTTCTGTGCGGTACCGGCG 742
QY 422 CCGGTTCTGTGCGGTGACCTTCTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481
DB 743 CACTGCGCTGTGCGGACCGGACTGTGCGAGGCGGCTTCCACTACTGTGCGGAGTGTGCG 802
QY 482 GGCAGACACCCAGAGGAGAGGAGGAGCCCGGATCTGTGCGGACCCGAGAACTGAGC 541
DB 803 TCGCCCGGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 862
QY 542 CGAGAGAGGTGG--CGGCGAGGAAAGCGGTCCCGGACCGGAGGAGGAGGAGGAGTGT 600
DB 863 CGGCGGCGGT 922
QY 601 CAGGCTGACGAGGAGGCGCGGAGTGTGAGCAGACGACGACGAGAGACCGGAACTGTG 660
DB 923 TCGACGTCTGTGCGGACCGGCTGCGCATGCGGCGCGGCGGCTGTGTGCGGTCTGACC 982
QY 661 GCGGCGATGAGTGTGAGCTGTGAGCGGAGG 691
DB 983 CCGGCTGT 1013

Search completed: April 6, 2004, 20:47:15
Job time : 538 secs

FT exon 5255..5324
 FT /*tag= f
 FT /number= 3
 FT 5325..6206
 FT /*tag= g
 FT /number= 3
 FT 6207..6352
 FT /*tag= h
 FT /number= 4
 PN WO20028314-A2.
 PD 07-NOV-2002.
 XX 26-APR-2002; 2002WO-US013245.
 PR 27-APR-2001; 2001US-0084864.
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 PA (AMHP) WEITH.
 PI Marzuk NM, Wang P, Bai Y, Wu X;
 XX WPI; 2003-167110/16.
 DR P-PSDB; ABP58235.
 PS Claim 1; Page 130-134; 141p; English.
 XX The present sequence is that of the murine ovary-specific OI-180 gene on
 CC chromosome 5. OI-180 clones were initially identified in a cDNA
 CC subtractive hybridisation screening using ovaries from Gd19 knockout mice
 CC and wild-type mice. OI-180 cDNA (see AB224589) was then used to isolate
 CC the OI-180 gene from a genomic library generated from mouse 129/SvEv
 CC strain. Loss of OI-180 results in female infertility and subfertility.
 CC The invention provides ovary-specific and oocyte-specific murine and
 CC human OI-180, OI-184 and OI-236 polynucleotides and polypeptides. These
 CC genes and their protein products appear to relate to various cell
 CC proliferative or degenerative disorders, especially those involving
 CC ovarian tumours, such as germ line tumours and granulosa cell tumours, or
 CC infertility, such as premature ovarian failure. The invention provides a
 CC method for detection of a cell proliferative or degenerative disorder of
 CC the ovary, which is associated with the expression of OI-180, OI-184 or
 CC OI-236. It also provides a method for treating such disorders by using an
 CC agent which suppresses or enhances the respective activities of OI-180,
 CC OI-184 or OI-236, and a method of screening for compounds that interact
 CC and/or modulate the expression or activity of the ovary-specific genes.
 CC These compounds are possible contraceptive or fertility enhancing agents.
 CC The modulator is preferably a polypeptide, small molecule or
 CC polynucleotide sequence
 XX
 XX Sequence 6873 BP; 1716 A; 1624 C; 1697 G; 1836 T; 0 U; 0 Other;
 SQ
 Query Match 62.6%; Score 799.4; DB 7; Length 6873;
 Best Local Similarity 99.9%; Pred. No. 8.1e-180;
 Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 2693 CAGCCACAGGCGGACAGCTCATGCGCTCTGTCGGATGAGTCCCGGTGCTAG 2752
 Qy 243 CAGCCGTAGCGTGGGTGAGAGTGAACCGGCGCGAGGCGCTCGGTGAGTTCAC 302
 Db 2753 CAGCCGTAGCGTGGGTGAGAGTGAACCGGCGCGAGGCGCTCGGTGAGTTCAC 2812
 Qy 303 CCGGCGCGGACAGCTGACGCTGACAGGATGCGAGGACCGCCGACGCGGATGCGGTT 362
 Db 2813 CCGGCGCGGACAGCTGACGCTGACAGGATGCGAGGACCGCCGACGCGGATGCGGTT 2872
 Qy 363 CTGTCAACCGGCGGACAGGCGGCGCGGAGATGCCCGGATCTCGGACACCGTATG 422
 Db 2873 CTGTCAACCGGCGGACAGGCGGCGCGGAGATGCCCGGATCTCGGACACCGTATG 2932
 Qy 423 CCGGTTCTGTCTGCTGACCTTCTGTGCGCTCTCTCTCTCACTGAGGTTGCGGAGGAC 482
 Db 2933 CCGGTTCTGTCTGCTGACCTTCTGTGCGCTCTCTCTCTCACTGAGGTTGCGGAGGAC 2992
 Qy 483 GCAGACACCCACGAAAGGAGAGGAGAGTCTCTCGGAGACCGGGAACCGGAGCC 542
 Db 2993 GCAGACACCCACGAAAGGAGAGGAGAGTCTCTCGGAGACCGGGAACCGGAGCC 3052
 Qy 543 GAGAGAGGTGCGCGGAGAAAGCGTCCCGCAGCGGAGAGCGAGAGGCGATGTTCA 602
 Db 3053 GAGAGAGGTGCGCGGAGAAAGCGTCCCGCAGCGGAGAGCGAGAGGCGATGTTCA 3112
 Qy 603 GCGTGCAGGAGAGGCGCGGATGAGAGCAGACACCGAGAGACCGGAACAGTGTGCG 662
 Db 3113 GCGTGCAGGAGAGGCGCGGATGAGAGCAGACACCGAGAGACCGGAACAGTGTGCG 3172
 Qy 663 GCGATGAGCTGTGAGCTGTGAGCGAGAGCAGATGCTCTCCGAGAGATGCTCAAGA 722
 Db 3173 GCGATGAGCTGTGAGCTGTGAGCGAGAGCAGATGCTCTCCGAGAGATGCTCAAGA 3232
 Qy 723 CCGCGTATGATGATGCGCTCTGAGAGCAGAGCTCTCCCGCAAGCAGGACGACAA 782
 Db 3233 CCGCGTATGATGATGCGCTCTGAGAGCAGAGCTCTCCCGCAAGCAGGACGACAA 3292
 Qy 783 GAGGCGCTGCGTTTCCAGTT 803
 Db 3293 GAGGCGCTGCGTTTCCAGTT 3313
 RESULT 4
 AB224591
 ID AB224591 standard; DNA; 4090 BP.
 AC AB224591;
 XX
 DT 31-MAR-2003 (first entry)
 XX
 DE Mouse ovary-specific OI-180 pseudogene.
 XX
 KM Ovary; OI-180; mouse; contraceptive; antiinfertility; cyostatic;
 KM gene therapy; pseudogene; ss.
 XX
 OS Mus musculus.
 XX
 PH Key Location/Qualifiers
 FT misc_feature 1..4090
 FT /*tag= a
 FT /note= "N represents unknown base"
 FT 1..786
 FT /*tag= b
 FT /number= 1
 FT 787..2616
 FT /*tag= c
 FT /number= 1
 FT 2617..2709
 FT /*tag= d
 FT /number= 2
 FT 2710..2817
 FT intron

PS Claim 2; Page 136-137; 141pp; English.

XX The present sequence is that of a human ovary-specific O1-180
CC polynucleotide. In mice, loss of O1-180 is associated with female
CC infertility and subfertility. The invention provides ovary-specific and
CC oocyte-specific murine and human O1-180, O1-184 and O1-236
CC polynucleotides and polypeptides. These genes and their protein products
CC appear to relate to various cell proliferative or degenerative disorders,
CC especially those involving ovarian tumours, such as germ line tumours and
CC granulosa cell tumours, or infertility, such as premature ovarian
CC failure. The invention provides a method for detection of a cell
CC proliferative or degenerative disorder of the ovary, which is associated
CC with the expression of O1-180, O1-184 or O1-236. It also provides a
CC method for treating such disorders by using an agent which suppresses or
CC enhances the respective activities of O1-180, O1-184 or O1-236, and a
CC method of screening for compounds that interact and/or modulate the
CC expression or activity of the ovary-specific genes. These compounds are
CC possible contraceptive or fertility enhancing agents. The modulator is
CC preferably a polypeptide, small molecule or polynucleotide sequence
XX

SC Sequence 2075 BP; 588 A; 413 C; 473 G; 601 T; 0 U; 0 Other;

Query Match 10.5%; Score 134.6; DB 7; Length 2075;
Best Local Similarity 71.1%; Pred. No. 6.6e-22;
Matches 207; Conservative 0; Mismatches 79; Indels 5; Gaps 2;

OY 968 CAAAGTTGTAAGAAGAACTAGATGTGCTGCCAGTCAAGATTGCGACGTGAACCTTAA 1027
DB 1360 CAGAGTTGTAAGAAGAACTAGATGTGCTGCCAGTCAAGATTGCGACGTGAACCTTAA 1419
OY 1028 CGCCCCCATGGGCAAGATTTGTGGGAGATGAGAGCAAGAACGCTGCTGGAGAGC 1087
DB 1420 CGCCCCCATGGGCAAGATTTGTGGGAGATGAGAGCAAGAACGCTGCTGGAGAGC 1479
OY 1088 ACCTTACAGCTTCAATATCATATTA-CTGAGAGTGCMAAACGTTTGTCTAGATGGGC 1146
DB 1480 ACTTACAGCTTCAATATCATATTA-CTGAGAGTGCMAAACGTTTGTCTAGATGGGC 1539
OY 1147 TAATGAGATGACAAAGTATGCTTCTCCCTCTCAGCTTCTCCCTTCCAAATTTCTCA 1206
DB 1540 GAGTGAAGAGATGAGCTTTTCGTCGCTCTCCCTCTCAGCTTCTCCCTTCCAAATTTCTCA 1599
OY 1207 TGACAGACAGTGTACTTGGATATTAAGCTGTGAATTAAGGATTTGCA 1257
DB 1600 TGAAGGACAGTGTATTCG---AAAAAGCTTCAATTAAGGATTTGCA 1646

RESULT 6
AAK53491
ID AAK53491 standard; DNA; 114955 BP.

AC AAK53491;
XX
XX 05-JUN-1999 (first entry)
XX
XX Human adenosine A1 receptor antisense oligonucleotide fragment.
DE Antisense oligonucleotide; multiple target; antisense treatment;
XX impaired respiration; inflammation; lung disease;
XX pulmonary vasoconstriction; inflammation; allergic rhinitis;
XX acute asthma; allergy; asthma; impeded respiration;
XX respiratory distress syndrome; pain; cystic fibrosis;
XX pulmonary hypertension; pulmonary vasoconstriction; emphysema;
XX chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
XX colon cancer; breast cancer; lung cancer; pancreatic cancer;
XX hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
XX prostate cancer; ss.
XX
XX Synthetic.
OS
XX WO913886-A1.
XX
XX PD 25-MAR-1999.

XX
XX 17-SEP-1998; 98WO-US019419.
XX
XX 17-SEP-1997; 97US-0059160P.
XX
XX 09-JUN-1996; 96US-0003972.
XX
XX (UYEC-) UNIV EAST CAROLINA.
XX
XX Nyce JW;
XX
XX WPI; 1999-229400/19.
XX
XX New antisense oligonucleotides used in treatment of, e.g. pulmonary
XX PT vasoconstriction.
XX
XX Disclosure; Page 37; 120pp; English.

XX The specification describes antisense oligonucleotides (AAK52869-X55271)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of RNAs corresponding to target genes, gene initiation
CC codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-
CC end and the juxta-section between coding and non-coding regions and all
CC segments of RNAs encoding proteins associated with one or more diseases,
CC conditions or mixtures. The antisense oligonucleotides may be derived
CC from sequences AAK5572-74. These multiple target oligonucleotides
CC (specifically AAK55180-271) can be used for the antisense treatment of
CC diseases and conditions. Typical diseases and conditions are those
CC associated with impaired respiration and inflammation, including lung
CC diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,
CC acute asthma, allergies, asthma, impeded respiration, respiratory
CC distress syndrome, pain, cystic fibrosis, pulmonary hypertension,
CC pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary
CC disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.
CC colon cancer, breast cancer, lung cancer, pancreatic cancer,
CC hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as
CC well as all types of cancers which may metastasize or have metastasized
CC to the lungs, including breast and prostate cancer
XX

XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;

Query Match 4.6%; Score 58.8; DB 2; Length 114955;
Best Local Similarity 32.6%; Pred. No. 0.0025;
Matches 251; Conservative 71; Mismatches 441; Indels 8; Gaps 3;

OY 3 GCGCGGCGAGCGCGGAGCGACACCATTTCCCGGAGAGAGTTCACACCTTCCGCCA 62
DB 104641 GCGCGGCGAGCGCGGAGCGACACCATTTCCCGGAGAGAGTTCACACCTTCCGCCA 104700
OY 63 TCTTATTCGCGAGGCCACCAAGACCGGGAGTGTGAGAGTTCGAGACCGCAGGGGTGCG 122
DB 104701 GC---GGCGSNNNDNNCCGCGGCGCGGCGGCGSNNNDNNCCGCGGCGGCGGCGG 104756
OY 123 ACCCGGCGCGCGCTCTCTCTCCCGGCTACAGACGCTACGCGCGGAGTACGTCA 182
DB 104757 NDNCCGCGGCGCGGCGGCGSNNNDNNCCGCGGCGGCGGCGGCGGCGGCGGCGG 104816
OY 183 CAGCCACAGCGCGGAGCAAGCTTATGCGCTGTGCGGATGGATCCCGGTGGTATG 242
DB 104817 SNNNDNNCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104876
OY 243 CAGCGGTAGCGTGGGAGTGAACCGCGGCGCGG--GAGCGCTCGATGAGTGTCA 300
DB 104877 NCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104936
OY 301 CTGCGCGCGCGGAGCGCTGAGCTGAGAGTGCAGGCGGCGGCGGCGGCGGCGGCGG 360
DB 104937 CCBGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 104996
OY 361 TCTGTCAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 104997 GSNNDNNBGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105056
OY 421 GCCCGCTTCTGTCGATGACCTTCTGTGGAGCTTCTCTCTCACTGAGAGTTGCGGAGAGC 480

[illegible]

	Query Match	4.5k; Score 57.8; DB 7; Length 2000;
	Best Local Similarity 10.4k; Pred. No. 0.0012;	
	Matches 79; Conservative 359; Mismatches 315; Indels 7; Gaps 4;	
QY	519 CTCGGGAGACCCGGAGACCGGAGCCCGAGAGAGTGCGCCCG-AGAGAAAGCGGTCCCCAGC 577	
DB	12 CSMRGRRRRYYTAAAGMMSCAAGSSRMSRKMGSXYKXKCSGCKXKTTTRKSRMYASS 71	
QY	578 CGGCAACCGAGAGAGGGGATGTTCAAGGCTCAGAGGACGCGCGGTGGAGACAGCAGCAC 637	
DB	72 ASGRITGKMSGSGSYSGMGKXKTKRBSRMTGRGRFRGMRBRMRMTMGSYRRCARSGRVAAG 131	
QY	638 CACCGAGAGACCGGAAACAGTGTGCGCGCGAT-GCAGTCTAGCCTGTGGAGCGAGACCA 696	
DB	132 SGRMMWGGKSRMSYMMWCYARGCSCKRKKSGSGWGTCKRRGARGGSGWSGATYKSGSM 191	
QY	697 TGTCTTCGCGGAGAGATGATGCTCAAGAACCCCGGTGATTCCGATGCGCCCTCGAGACCAAGCC 756	
DB	192 SKRMMMSCGSGCGRRSAYSRYTGSRKTYTKXMTYBASRMRMYMTTSYMASSY 251	
QY	757 TCCCGGAAAACACGAGACGAGACCAAGAGACGCGCTCGTTCCAGTTCTTAGACCAAG 816	
DB	252 TWCSRKRSMWKKWRRMRMRMSRSGWMSYKMMCTAYKKSYSFMCYWRGGGAGAT 311	
QY	817 TAOCGGTACATCATCTGCAGAGACTGCAGAAATCCGGTGGAGAGCGGCTATGTGTGGTGT 876	
DB	312 RYWRGWSRRAAMMYTKMTYRGYKMRGWMAGRMMRMRMRMRMRMRMRMRMRMRMRMR 371	
QY	877 GTGACGAGGACCAAGTAAAGTGTATTCAATCAACAGTTTGCGAGTGTGAGAAATCCTTA 936	
DB	372 RRRWAKKSRTSRKX 429	
QY	937 CAACCTTACAGATGAGAGACATCACCTGTAAGAGTTGTAAGAAAGAACTGATGTGCGTG 996	
DB	430 GCMTCRMKSYGMKMKWSKMKWASKTKMMSMTMRKKKCSRTTWMGTRGSMGT---MG 486	
QY	997 CCCAGTCAGATTTCGCACCGTGAACCCCTAAACGCCCCCATCGGCAAGACTTGTGTGGAG 1056	
DB	487 RCRYKXKSGMKRKCRRBRMGMYMRMYMRKYMSARMTYMRCAKXKYSVSAARKRACWYRG 546	
QY	1057 ATGCAAGACCAAGCGCTGTCCTGCGACAGACACTTGAAGCTTCAATAACATCATTTAGTG 1116	
DB	547 KGYWAGKMMRMYRMYYKMMWYKXKYSKCSWYCKMSYTAACMSKARAGAKCMRCKSK 606	
QY	1117 AGAGTCGAAAACGTTTGTCTAGATGGGGCTAATGAATGACAAGTAGAGCTTCTCCCC 1176	
DB	607 MSASMSMRSRKRCRKCASKRSSAKRYAMMGMTSGSRMSRMYKTCYWRKMGSMKSRCT 666	
QY	1177 TCTTCACCTCTCCCTTCCAAATTTCTCATGACAGACAGTGTAAGTACTTGATATATAAGCC 1236	
DB	667 WMYYSKTYTAKYKSYMYRYPRAACMTWRYRYRYSYMTYMYATYSITMAATGMKIS 726	
QY	1237 TGTGAATTAAGATATTGCAAAACAAAAAATAAAAA 1275	
DB	727 GRWTSWYKXCKSGMKRYRSMWYYSMMWAKTWKMBRYA 765	
RESULT 8		
ADA69587		
ID	ADA69587 standard; DNA; 1032 BP.	
XX		
AC	ADA69587;	
XX		
DT	20-NOV-2003 (first entry)	
XX		
DE	Rice gene, SEO ID 2910.	
XX		
KW	Plant; bacterial infection; fungal infection; viral infection; rice;	
XX	gene; ds.	
OS	Oryza sativa.	
XX		

PN WO2003000898-A1.
 XX 03-JAN-2003.
 XX 22-JUN-2001; 2001MO-IB001105.
 XX 22-JUN-2001; 2001MO-IB001105.
 XX 22-JUN-2001; 2001MO-IB001105.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 XX Katerberg F, Quan S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 XX Identifying at least one gene involved in plant resistance or response to
 XX pathogenic infection for conferring resistance or tolerance to a plant to
 XX bacterial, fungal or viral infection by determining or detecting plant
 XX gene expression.
 XX Claim 6; SEQ ID NO 2910; 899pp; English.
 XX The present invention relates to a method (M1) for identifying genes
 XX involved in plant resistance or response to pathogenic infection. M1
 XX comprises identifying a gene whose expression is significantly altered in
 XX the incompatible interaction of plant gene expression relative to
 XX expression of the gene in an uninfected plant, in a mutant plant that
 XX does not express a gene associated with response to pathogenic infection,
 XX or in a corresponding incompatible or compatible interaction. (M1) is
 XX useful for conferring resistance to resistance or tolerance to a plant to
 XX bacterial, fungal or viral infection. The present sequence was used to
 XX illustrate the invention.
 XX Sequence 1032 BP; 189 A; 331 C; 394 G; 118 T; 0 U; 0 Other;
 SQ
 Query Match 4.2%; Score 54.2; DB 7; Length 1032;
 Best Local Similarity 43.4%; Pred. No. 0.007;
 Matches 251, Conservative 0; Mismatches 328; Indels 0; Gaps 0;
 QY 137 CTTTCCTCCCGGCTAGACAGCTCATGCGCCGCGAGTACGTCACACCCACGCGG 196
 Db 167 CTTTCCTCCCGGCTAGACAGCTCATGCGCCGCGAGTACGTCACACCCACGCGG 226
 QY 197 CACAGCTATGCGCGCTGCTGCGCGATGATCCCGGTCGTCACAGCCGTCGCTG 256
 Db 227 AGAAGAGCGCGCTGACCGCGAGAGTGCATGCTGCGAGAGAGCTTCGAGGAG 286
 QY 257 CGGTGAGTGAAACCGCGCGCGCGAGCTGCGAGTGTCTACTCGGCGCGCGACCG 316
 Db 287 AGAACAAGCTGAGACCGCGAGCGAGAGAGCTGCGCGAGTGTCTACTCGGCGCGCGACCG 346
 QY 317 TGACAGCTGAGAGGTGCGGAGCGAGCGCGCGAGTGTCTGTCACACCCGTCG 376
 Db 347 GCGAGTGTGCGGTGTGTGTCAGAAACCGCGCGCGCTGAGAAAGACCAAGCTGAGAC 406
 QY 377 GCGACGCGCGCGCGCGAGATCCCGCGATCTGTGAGACCGTACCGCGCGCTGTCGCG 436
 Db 407 GCGACTTGCACCGCGCTCAAGCGCGCTGTGACCGCGCGCGCGCGACCGACCGCTCC 466
 QY 437 TGACCTTGTGCGCTTCTCTCTCACTGAGAGTGTGCGGAGGCGAGGACCGACCGACGA 496
 Db 467 TCCAGGAAACACCGCGCTCTCACTGAGTGTGTGTTGACCGAAAGCTGCAAGAGA 526
 QY 497 AGGAGAGGAGGAGCGCGGATCTCGGAGACCGCGGAAACCGAGCCGAGAGAGTGGCGG 556
 Db 527 AGGAGAGAGAGCGAGGAGGAGCGCGCGCGCGCTGACGTCGCGCGCTTCTGCGCG 586
 QY 557 CGAAGAAAGCGGTCCCGACCGCGGAAACCGAGAGGCGGATTTCCAGCTGCGAGGCGAG 616
 Db 587 CGGCGAGTGAAGTGTGCGGTCCCGGACCGCGAGAAACCGGCGCTGAGAGAGCGCGCG 646
 QY 617 CCGGTGAGAGAGACAGACCAACCGAGAGACCGGAAACAGTGTGCGCGGATGCACTCTG 676

Db 647 CGGCTTCGAGAGACAGACAGACGAGTGAAGCCCGAGAGAGGCTGAGACGCGCA 706
 QY 677 AGCCTGGAGAGAGACCATGTCCTGCGCGAGAGATGG 715
 Db 707 CGGCGGAGAGCGCGTGTGTGAGACGAGCGCGACTGG 745
 RESULT 9
 ID AAX53491 standard; DNA; 114955 BP.
 XX AAX53491;
 XX 05-JUN-1999 (first entry)
 XX Human adenosine A1 receptor antisense oligonucleotide fragment.
 XX Antisense oligonucleotide; multiple target; antisense treatment;
 XX impaired respiration; inflammation; lung disease;
 XX pulmonary vasoconstriction; inflammation; allergic rhinitis;
 XX acute asthma; allergy; asthma; impeded respiration;
 XX respiratory distress syndrome; pain; cystic fibrosis;
 XX pulmonary hypertension; pulmonary vasoconstriction; emphysema;
 XX chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
 XX colon cancer; breast cancer; lung cancer; pancreatic cancer;
 XX hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
 XX prostate cancer; ss.
 XX Synthetic.
 XX WO9913886-A1.
 XX 25-MAR-1999.
 XX 17-SEP-1998; 98WO-US019419.
 XX 17-SEP-1997; 97US-0059160P.
 XX 09-JUN-1998; 98US-00093972.
 XX (UNEC-) UNIV EAST CAROLINA.
 XX Nyce JW;
 XX WPI; 1999-223400/19.
 XX New antisense oligonucleotides used in treatment of, e.g. pulmonary
 XX vasoconstriction.
 XX Disclosure; Page 37; 120pp; English.
 XX The specification describes antisense oligonucleotides (AAX52869-X55271)
 XX directed against at least 2 mRNAs selected from target genes, coding and
 XX non-coding regions of RNAs corresponding to target genes, gene initiation
 XX codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-
 XX end and the juxta-region between coding and non-coding regions and all
 XX segments of RNAs encoding proteins associated with one or more diseases,
 XX conditions or mixtures. The antisense oligonucleotides may be derived
 XX from sequences AAX55180-271. These multiple target oligonucleotides
 XX (specifically AAX55180-271) can be used for the antisense treatment of
 XX diseases and conditions. Typical diseases and conditions are those
 XX associated with impaired respiration and inflammation, including lung
 XX diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,
 XX acute asthma, allergies, asthma, impeded respiration, respiratory
 XX distress syndrome, pain, cystic fibrosis, pulmonary hypertension,
 XX pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary
 XX disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.
 XX colon cancer, breast cancer, lung cancer, pancreatic cancer,
 XX hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as
 XX well as all types of cancers which may metastasize or have metastasized
 XX to the lungs, including breast and prostate cancer
 XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;
 SQ

XX 17-FEB-1999; 99WO-US003265.
 XX 18-FEB-1998; 98US-00024753.
 PR 05-MAY-1998; 98US-00072596.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R,
 PI Vedvick TS, Twardzik DR, Lodes MT, Hendrickson RC,
 XX WPI; 1999-527416/44.
 DR
 XX
 XX
 PT New polypeptide comprising antigenic portions of M. tuberculosis.
 XX
 PS Claim 4; Page 215-216; 323pp; English.
 XX
 CC This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against M.
 CC tuberculosis infection. The new detection methods are needed as current
 CC vaccination strategies do not provide 100% immunity
 XX
 SQ Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 U; 0 Other;

Query Match 3.9%; Score 49.6; DB 2; Length 400;
 Best Local Similarity 49.2%; Pred. No. 0.064;
 Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY	458	CCCTCACTGAGAGGTTGGGGAGGAGGACACCAACGAAAGGAGAGGGAGCCCGGCAT	517
DB	119	CTTCAACCCAAAGGCGCCGACGACGACCGCCGAGTGAACGCGGGTGGCGGCA	178
QY	518	CCTCGGGGACCCCGGAAACCGGACCGAGAGGTGCGCGAGAAAGCGTCCCCAGC	577
DB	179	CGCGGGAACGGCGGAAACGGCGGAGACCAACACACCGCGCGCGCGGACCAAGG	238
QY	578	CGGAAAGCGAGAGGGGATGTTCAAGCTGCAAGGCGAGGCGGGTGGAGCAGCCAC	637
DB	239	CGGCGACGGGGGGCGCGCGGGCGGAAACGGGAAACCGGCGAGCGCGCGC	298
QY	638	CACCGAGGACCGGAAACAGTGTGCGGCGGATGCACTGAGCCTGGAGCGAGAGCCAT	697
DB	299	CGGACACCGGGGCGGACCAACGCAACCGCGGCAACCGCGGCAACGCGGCAAGGCGC	358
QY	698	GTCCTGCCGCAAGATGGCTCAGG	721
DB	359	CGGCGGCGACGTTGCACTTCAGG	382

Search completed: April 6, 2004, 18:08:48
 Job time : 622 secs